

Db 440 CTGCGCTGCTGAGGCTTACGCGGCTGCGCCCGCAAGAGATGAGCTCGGCACT 439
Qy 421 GAGGTCTGAGCTGATACCAAGCTCTCTACGCTTTTCTCTCGAGCCATGAGGACATG 480
Db 500 GAGGTCTGAGCTGATACCAAGCTCTCTACGCTTTTCTCTCGAGCCATGAGGACATG 559
Qy 481 CACAGCTGATGAGAGCGGCTCACCGTATCTCTGAGCTTACGCTTCTTCTGCG 540
Db 560 CACAGCTGATGAGAGCGGCTCACCGTATCTCTGAGCTTACGCTTCTTCTGCG 619
Qy 541 CAGATGAGCAAGCGGCTGAGGCACTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 620 CAGATGAGCAAGCGGCTGAGGCACTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 679
Qy 601 CTGATGAGCTTGTCTTCTGCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 680 CTGATGAGCTTGTCTTCTGCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
Qy 661 GAGCTCTGAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 740 GAGCTCTGAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 799
Qy 721 TACGTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 800 TACGTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 859
Qy 781 TGAAGCTGAG 840
Db 860 TGAAGCTGAG 919
Qy 841 GAGCTCTGAGCTTGTCTTCTGCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 920 GAGCTCTGAGCTTGTCTTCTGCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 979
Qy 901 TGGAGCTGAG 960
Db 980 TGGAGCTGAG 1039
Qy 961 ACAGGCAAG 1006
Db 1040 ACAGGCAAG 1085

RESULT 2
US-09-220-132-135/c
Sequence 135, Application US/09220132

GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220,132
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25 US 60/068,861
PRIOR APPLICATION NUMBER: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 135
LENGTH: 2559
TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-135

Query Match 24.5%; Score 505.2; DB 4; Length 2559;
Best Local Similarity 94.6%; Pred. No. 1,4e-99;
Matches 566; Conservative 0; Mismatches 28; Indels 4; Gaps 4;

Qy 604 TGTGCTTTGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663

Db 1634 TGAACCTTCTGTTTCTCCCAATGTCACAGAAAGAGCTGATGAGAGAGAGAGAG 1575
Qy 664 TCTGCTGCTGATGAG 723
Db 1574 TCTGCTGCTGATGAG 1516
Qy 724 GTGGAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 783
Db 1515 GTGGAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1456
Qy 784 AGCTTGGAG 843
Db 1455 AGCTTGGAG 1396
Qy 844 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
Db 1395 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1336
Qy 904 GCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
Db 1335 GCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1276
Qy 964 GCGCAAG 1023
Db 1275 G-CACAG 1218
Qy 1024 GATCCATCTGAG 1083
Db 1217 GATCCATCTGAG 1159
Qy 1084 AGGAG 1143
Db 1158 AGGAG 1099
Qy 1144 TCACTGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1201
Db 1098 TCACTGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1041

RESULT 3
US-08-146-421-4/c
Sequence 4, Application US/08146421

GENERAL INFORMATION:
APPLICANT: BREWER, GARY
TITLE OF INVENTION: DNA SEQUENCE ENCODING A POLYPEPTIDE WITH
TITLE OF INVENTION: ANTI-TUMOR PROPERTIES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DILMORTH & BARRESE
STREET: 4350 LA JOLLA VILLAGE DRIVE, SUITE 300
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/146,421
FILING DATE: 29-OCT-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: PEPPER PH.D., FREDERICK W.
REGISTRATION NUMBER: 31,286
REFERENCE/DOCKET NUMBER: 489-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4410
TELEFAX: 619-453-2839
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

121 CCCAGACTGCCCCCTGCTGTTGCCCCCTGAGCCCACTATGCTCCCAAGATGAGCACT 180
181 GCTGTGGCACTGCTCCCTGCTGTTGAGCCCTATGCTCTGAGAGCCGAGAGAGGCGG 240
181 GCTGTGGCACTGCTCCCTGCTGTTGAGCCCTATGCTCTGAGAGCCGAGAGAGGCGG 240
241 CGGAGCTTACCGAGGCTCTGCACTGCTGCTTACAGAGCACTGAGTATACCTGCAAGTATCC 300
241 CGGAGCTTACCGAGGCTCTGCACTGCTGCTTACAGAGCACTGAGTATACCTGCAAGTATCC 300
301 GTCCAGAGAGCCCTGAGGCTGCTGAGGCTTACAGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
301 GTCCAGAGAGCCCTGAGGCTGCTGAGGCTTACAGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
361 GCTGAGGCTTACCGAGGCTCTGCACTGCTGCTTACAGAGCACTGAGTATACCTGCAAGTATCC 420
361 GCTGAGGCTTACCGAGGCTCTGCACTGCTGCTTACAGAGCACTGAGTATACCTGCAAGTATCC 420
421 CATGAGGCACTGCACTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 480
421 CATGAGGCACTGCACTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 480
481 GTGCTCTTCCGCGCAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 540
481 GTGCTCTTCCGCGCAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 540
541 CTGATATCTCAAGCTGT 600
541 CTGATATCTCAAGCTGT 600
601 GAGAACTTGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
601 GAGAACTTGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
661 GCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 720
661 GCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 720
721 GCAGCCGATGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 780
721 GCAGCCGATGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 780
781 TTCCAGAGCTTGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
781 TTCCAGAGCTTGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
841 CTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 900
841 CTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 900
901 GCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 960
901 GCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 960
961 CCCTTACCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1020
961 CCCTTACCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1020
1021 GAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1074
1021 GAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1074

RESULT 2
US-09-799-875-7
Sequence 7, Application US/09799875
Patent No. US20020034780A1
GENERAL INFORMATION:
APPLICANT: Meyer, Rachel
APPLICANT: Kapeller-Lidemann, Rosana
APPLICANT: Williamson, Mark
TITLE OF INVENTION: No. US20020034780A1 Human Protein Kinases and Uses

TITLE OF INVENTION: Therefor
FILE REFERENCE: 35800/20996
CURRENT APPLICATION NUMBER: US/09/799, 875
CURRENT FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: 60/182, 059
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: 09/559, 287
PRIOR FILING DATE: 2000-09-12
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2389
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (383)....(1456)
US-09-799-875-7

Query Match 99.78; Score 1070.8; DB 9; Length 2389;
Best Local Similarly 99.88; Pred. No. 3.9e-272;
Matches 1072; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGCGAGCCACCCCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGG 60
383 ATGCGAGCCACCCCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGG 442
61 TTGATGACAACTTGAATACCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 120
443 TTGATGACAACTTGAATACCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 502
121 CCCAGACTGCCCCCTGCTGTTGAGCCCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 180
503 CCCAGACTGCCCCCTGCTGTTGAGCCCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 562
181 GCTGTGGCACTGCTCCCTGCTGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 240
563 GCTGTGGCACTGCTCCCTGCTGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 622
241 CGGAGCTTACCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 300
623 CGGAGCTTACCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 682
301 GTCCAGAGGCTTGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
683 GTCCAGAGGCTTGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
361 GCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 420
743 GCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 802
421 CATGAGGCACTGCACTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 480
803 CATGAGGCACTGCACTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 862
481 GTGCTCTTCCGCGCAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 540
863 GTGCTCTTCCGCGCAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 922
541 CTGATATCTCAAGCTGT 600
923 CTGATATCTCAAGCTGT 982
601 GAGAACTTGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
983 GAGAACTTGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1042
661 GCTGAGGCTTACCGAGGCTCTGCACTGCTGCTTACAGAGCACTGAGTATACCTGCAAGTATCC 720
1043 GCTGAGGCTTACCGAGGCTCTGCACTGCTGCTTACAGAGCACTGAGTATACCTGCAAGTATCC 1102
721 GCAGCCGATGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 780

Accession	Sequence	Position
Dh	GCAGCCGATGCTGTGAGAGCTGGGCGGTGGCTTTACCATGCTGGCGGACCATACCCC	1162
Qy	TTCCAGCATCTGGAGCCCTCTCTCTTTGGGCAAGATCCGCGCGGGGCGCTTACGCGTTG	840
Dh	TTCCAGGACTGTGGAGCCCTGTCTCTGCTCTTTCGGCAAGATCCGCGCGGGGCGCTTACGCGTTG	1222
Qy	CCTGCAGGCGCTCTCGGCCCCCTGGCCGCGCTGTCTGGTTCTCGCTGCTCTTTCGTTCGGAGGCCA	900
Dh	CCTGCAGGCGCTCTCGGCCCCCTGGCCGCGCTGTCTGGTTCTCGCTGCTCTTTCGTTCGGAGGCCA	1282
Qy	GCTGAAACGGCTTCACAGCCACAGGATCCTCCTGCAACCCCTGGCTGGGACAGAACCCGATG	960
Dh	GCTGAACGGCTTCACAGCCACAGGATCCTCCTGCAACCCCTGGCTGGGACAGAACCCGATG	1342
Qy	CCCTTAGCGCCCAACCCGATTCGCATCTCTGGAGGCTGGCCAGAGTGTCTCTGATGACTG	1020
Dh	CCCTTAGCGCTTCACAGCCACAGGATCCTCCTGCAACCCCTGGCTGGGACAGAACCCGATG	1402
Qy	GGGCTGACAGAAAGCAGGAGAGAGAGGAGACAGAGAAAGTGTTCTGTATGGC	1074
Dh	GGGCTGACAGAAAGCAGGAGAGAGAGGAGACAGAGAAAGTGTTCTGTATGGC	1456

RESULT 3

```

Sequence 271.  Application US/10098841
Publication No. US20020197679A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundqing
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Mehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Drmanac, Radcoje T.
TITLE OF INVENTION: No. US20020197679A1 Nucleic Acids and
FILE REFERENCE: 784CIP2
CURRENT APPLICATION NUMBER: US/10/098,841
CURRENT FILING DATE: 2002-03-13
PRIORITY APPLICATION NUMBER: 09/599,042
PRIORITY FILING DATE: 2000-06-20
PRIORITY APPLICATION NUMBER: 09/552,317
PRIORITY FILING DATE: 2000-04-25
PRIORITY APPLICATION NUMBER: 09/488,725
PRIORITY FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 331
SOFTWARE: pc_FL_genes Version 1.0
SEQ ID NO 271
LENGTH: 2092
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
US-10-098-841-271

```

Query Match	99.6%	Score 1069.2	DB 14	Length 2032
Best Local Similarity	99.7%	Pred. No. 1e-271		
Matches 1071; Conservative	0	Mismatches 3	Indels 0	Gaps 0

Oy 1 ATGCGAGCACCCCTCTGGCTGCTCTGGGGTTCCTGTCCAGAAAGCCGTTGGAG 60
Db 139 ATGCGAGCACCCCTCTGGCTGCTCTGGGGTTCCTGTCCAGAAAGCCGTTGGAG 198

QY	61	TTGGAATGACAACCTTAGATATACCGAGCGTCCCGCTCAGAAACAGAGTTCGAAATGGGCCCCAG	120
Dp	199	TTGGATATACAACTTAGATATACCGAGCGTCCCGCTCAGAAACAGAGTTCGAAATGGGCCCCAG	258
QY	121	CCGAGACTGGCCCCCTGGCTGTGTGGCCCTGAGGCCACTATCTGTCTCAGATTCGTGCAACT	180
Dp	259	CCGAGACTGGCCCCCTGGCTGTGTGGCCCTGAGGCCACTATCTGTCTCAGATTCGTGCAACT	318
QY	181	GCTGTGGCCACTGGCTCTCCCGCTTTTGGGCGCTTAATGTCTCTCTGAGAGCCCGAGAGAGGGGGG	240
Dp	319	GCTGTGGCCACTGGCTCTCCCGCTTTTGGGCGCTTAATGTCTCTCTGAGAGCCCGAGAGAGGGGGG	378
QY	241	CGGGCTTAACCGGGGCCCTGCACTGGCCCTTAACAGGCACTAGATATACCTGCAAGGTGTACCCC	300
Dp	379	CGGGCTTAACAGGCGCCCTGCACTGGCCCTTAACAGGCACTAGATATACCTGCAAGGTGTACCCC	438
QY	301	GTCCAGGAGACCCCTGGCCGTGTGGAGCCCTTAACGCGGCGTGGCCCGGCAACAGCATGTG	360
Dp	439	GTCCAGGAGACCCCTGGCCGTGTGGAGCCCTTAATGCGGGCGTGGCCCGGCAACAGCATGTG	498
QY	361	GCTCGGGCCACTAGAGTCTCGGCTGTGTATCCACAGTCTCTTAACGCTTTTTCACCTCGGAGC	420
Dp	499	GCTCGGGCCACTAGAGTCTCGGCTGTGTATCCACAGTCTCTTAACGCTTTTTCACCTCGGAGC	558
QY	421	CATGGGGACATGCAACAGCCTGATGTCGAAAGCCGCAACCGTATCCCTGAGCCCTGAGGCTGCC	480
Dp	559	CATGGGGACATGCAACAGCCTGATGTCGAAAGCCGCAACCGTATCCCTGAGCCCTGAGGCTGCC	618
QY	481	GTGCTCTTCCGGCCAGATATGCGCACCGCCCTGGCGCACTGTCAACAGACCGGTCTGTGCTG	540
Dp	619	GTGCTCTTCCGGCCAGATATGCGCACCGCCCTGGCGCACTGTCAACAGACCGGTCTGTGCTG	678
QY	541	CGTGAATCTCAAGCTGTGTCCGCTTTGTGTTGGCTGACCGTGTAGAGGAAGAAAGCTGTGTGCTG	600
Dp	679	CGTGAATCTCAAGCTGTGTCCGCTTTGTGTTGGCTGACCGTGTAGAGGAAGAAAGCTGTGTGCTG	738
QY	601	GAGAACCTTGAGAGACTCTGCTGATCTGACTGAGGCGCAGATATTTCCCTGTGGAGCAAGAC	660
Dp	739	GAGAACCTTGAGAGACTCTGCTGATCTGACTGAGGCGCAGATATTTCCCTGTGGAGCAAGAC	798
QY	661	GCGTGGCCAGCCTTACGTGGAGCCTGAGATATCTACGCTCACGGGCTCATATCTCGGGCAAG	720
Dp	799	GCGTGGCCAGCCTTACGTGGAGCCTGAGATATCTACGCTCACGGGCTCATATCTCGGGCAAG	858
QY	721	GACGCCGATATCTGAGAGCCTGGGGCGTGGGCGCTTTCACATGCTGTGGCGGCATACACCC	780
Dp	859	GACGCCGATATCTGAGAGCCTGGGGCGTGGGCGCTTTCACATGCTGTGGCGGCATACACCC	918
QY	781	TTTCAGAGACTCGAGAGCTGTCTGTCTTTCGGGCAAGATTCGCGCGCGGGGACTACGCCCTTG	840
Dp	919	TTTCAGAGACTCGAGAGCTGTCTGTCTTTCGGGCAAGATTCGCGCGCGGGGACTACGCCCTTG	978
QY	841	CTTGACAGGCTCTTGCGGCCCTTGGCCCGGTGTCTGTGTTGTGCTGCTCTTTCGTGGAGACCA	900
Dp	979	CTTGACAGGCTCTTGCGGCCCTTGGCCCGGTGTCTGTGTTGTGCTGCTCTTTCGTGGAGACCA	1038
QY	901	GCTGAACGAGCTCAACGCAACGAGCATCTGTCTGCAACCCCTGGGCTGGGACAGAGACCGGATG	960
Dp	1039	GCTGAACGAGCTCAACGCAACGAGCATCTGTCTGCAACCCCTGGGCTGGGACAGAGACCGGATG	1098
QY	961	CCCTTAGCCCCCAACCCGATCCCATCTCTGAGAGGCTGGCCAGGTGTGCTCTGTATGTGACTG	1020
Dp	1099	CCCTTAGCCCCCAACCCGATCCCATCTCTGAGAGGCTGGCCAGGTGTGCTCTGTATGTGACTG	1158
QY	1021	GGGCTGTGACGAACCCAGGGAAGAGAGGAGACAAGAGAAATGTGTTCTGTATGTGC	1074
Dp	1159	GGGCTGTGACGAACCCAGGGAAGAGAGGAGACAAGAGAAATGTGTTCTGTATGTGC	1212

RESULT 4

US-10-024-828-3
; Sequence 3, Application US/10024828

212

```
Publication No. US20030036051A1
GENERAL INFORMATION:
APPLICANT: Vlica, Duke
APPLICANT: Bird, Timothy A.
APPLICANT: Anderson, Dirk M.
APPLICANT: Marken, John S.
TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase
TITLE OF INVENTION: Functions
FILE REFERENCE: 2877-US
CURRENT APPLICATION NUMBER: US/10/024, 828
CURRENT FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: US/09/509, 902A
PRIOR FILING DATE: 1999-08-03
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1085
TYPE: DNA
ORGANISM: Homo sapiens
US-10-024-828-3
```

```
Query Match 89.1%; Score 956.4; DB 15; Length 1085;
Best Local Similarity 99.9%; Pred. No. 4.7e-242;
Matches 957; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGCGAGCCACCCCTTGGCTGCTCTGCGGGTTCCCTGTCCAGGAAGACGGTTGGAG 60
DB 128 ATGCGAGCCACCCCTTGGCTGCTCTGCGGGTTCCCTGTCCAGGAAGACGGTTGGAG 187
QY 61 TTGGATGACAACCTTAGATACCGAGCGTCCCTCCAGAAACAGAGCTCGAAGTGGGCCAG 120
DB 188 TTGGATGACAACCTTAGATACCGAGCGTCCCTCCAGAAACAGAGCTCGAAGTGGGCCAG 247
QY 121 CCCGAGCTGCCCCCTGCTGTGCTGCTGAGCCCACTACTGCTCAGATCGTCAACT 180
DB 248 CCCGAGCTGCCCCCTGCTGTGCTGCTGAGCCCACTACTGCTCAGATCGTCAACT 307
QY 181 GCTGTGCGCACTGCTCCGCTTGGGCGCTATGCTCCCTGAGAGCCCGAGAGAGGCGGG 240
DB 308 GCTGTGCGCACTGCTCCGCTTGGGCGCTATGCTCCCTGAGAGCCCGAGAGAGGCGGG 367
QY 241 CGGCGCTACCGGGCCCTGCACTGCTGCTACAGAGCACTGAGTATACCTGCAAGTGTACCC 300
DB 368 CGGCGCTACCGGGCCCTGCACTGCTGCTACAGAGCACTGAGTATACCTGCAAGTGTACCC 427
QY 301 GTCCAGAGAGCCCTGCGCGCTGCTGAGCCCTACCGCGCGGCTGCCCCGACAGCATGTG 360
DB 428 GTCCAGAGAGCCCTGCGCGCTGCTGAGCCCTACCGCGCGGCTGCCCCGACAGCATGTG 487
QY 361 GCTGCGCCCACTGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 488 GCTGCGCCCACTGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547
QY 421 CATGCGGACATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 548 CATGCGGACATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
QY 481 GTGCTCTTCCGCGCAGATGAGCCACCGCTGCGCGCACTGTCAACAGACAGGTCTGCTGTG 540
DB 608 GTGCTCTTCCGCGCAGATGAGCCACCGCTGCGCGCACTGTCAACAGACAGGTCTGCTGTG 667
QY 541 CGTGAATCTCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 668 CGTGAATCTCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 727
QY 601 GAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 728 GAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 787
QY 661 GCGTGGCCCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 788 GCGTGGCCCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 847
```

```
QY 721 GCAGCCCATGCTCTGAGAGCTGAGGCGTGGCGCTCTTACACAGCTGAGCGGCACTAACCC 780
DB 848 GCAGCCCATGCTCTGAGAGCTGAGGCGTGGCGCTCTTACACAGCTGAGCGGCACTAACCC 907
QY 781 TTCCAGAGACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 908 TTCCAGAGACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 967
QY 841 CTTGAGGCTCTGCGGCCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 968 CTTGAGGCTCTGCGGCCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027
QY 901 GCTGAAGGCTCAAGAGCAAGAGGATCTCTGCAACCCCTGAGCGGCTGAGCAAGACCCGA 958
DB 1028 GCTGAAGGCTCAAGAGCAAGAGGATCTCTGCAACCCCTGAGCGGCTGAGCAAGACCCGA 1085
```

RESULT 5

```
US-09-925-301-260
Sequence 260, Application US/09925301
Patent No. US20020052308A1
```

```
GENERAL INFORMATION:
```

```
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
```

```
FILE REFERENCE: P0106
CURRENT APPLICATION NUMBER: US/09/925,301
```

```
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
```

```
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
```

```
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
```

```
SEQ ID NO 260
LENGTH: 2048
```

```
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURES:
```

```
NAME/KEY: misc_feature
LOCATION: (66)
```

```
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (67)
```

```
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-260
```

```
Query Match 81.9%; Score 879.2; DB 9; Length 2048;
Best Local Similarity 92.6%; Pred. No. 1e-221;
```

```
Matches 994; Conservative 3; Mismatches 1; Indels 76; Gaps 4;
```

```
QY 1 ATGCGAGCCACCCCTTGGCTGCTCTGCGGGTTCCCTGTCCAGGAAGACGGTTGGAG 60
DB 156 ATGCGAGCCACCCCTTGGCTGCTCTGCGGGTTCCCTGTCCAGGAAGACGGTTGGAG 215
QY 61 TTGGATGACAACCTTAGATACCGAGCGTCCCTCCAGAAACAGAGCTCGAAGTGGGCCAG 120
DB 216 TTGGATGACAACCTTAGATACCGAGCGTCCCTCCAGAAACAGAGCTCGAAGTGGGCCAG 275
QY 121 CCCGAGTGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 276 CCCGAGTGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304
QY 181 GCTGTGCGCACTGCTCCGCTTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 305 -----GAGCCGAGAGAGGCGGG 322
QY 241 CGGCGCTACCGGGCCCTGCACTGCGCTACAGAGCACTGATATACCTGCAAGGTGATACCC 300
DB 323 CGGCGCTACAGAG-CTGCACTGCGCTACAGAGCACTGATATACCTGCAAGGTGATACCC 380
QY 301 GTCCAGAGAGCCCTGCGCGCTGCTGAGCCCTACGCGCGGCTGCGCCCGCAAGACATGTG 360
DB 381 GTCCAGAGAGCCCTGCGCGCTGCTGAGCCCTACGCGCGGCTGCGCCCGCAAGACATGTG 440
```


US-10-291-172-116

Query Match 24.5%; Score 263.2; DB 12; Length 1909;
 Best Local Similarity 60.7%; Pred. No. 1,2e-59;
 Matches 447; Conservative 0; Mismatches 288; Indels 1; Gaps 1;

```

QY 210 CTATGCTCTCTGAGAGCCGAGAGAGGCGGCGGCTACCGGCGCCCTGACCTGACCCCTAC 269
DB 103 CTACCTGCTGCTGCGCCCTAGCCGAGCGAGCATGTGTCCCGGCGCTGTGACATCCACAC 162
QY 270 AGGCACTGAGTATAC-CTGCAGAGTGTACCCCTCCAGAGAGCCCTGCGCTGCTGAGAC 328
DB 163 TGGAGCGCGGAGTGGCTGCAAGGTCTTCCATTAAACATACAGAGCAAAATCAGGC 222
QY 329 CCTACGCGGCGGCTGCGCCGACAGCATGTGCTGCGCCCACTGAGTCTGCTGCTGTA 388
DB 223 CTTACATCCAGCTGCGCATGCGACAGCAATTACTGCGCATTTGTGAGAGTGTCTTGGGG 282
QY 389 CCCAGCTCTTACGCGCTTTTTCATCTCGACCAATGGGACATGCAACGCTGTGTGCGAA 448
DB 283 AAACCAAGGCTATGTCTTCTTGTAGAGAGACTTTGGGACATGCACTCTATGTGCGAA 342
QY 449 GCCGCAACCGTATCCCTGAGCGCTGAGCGCTGCTCTTCCGCGAGATGCGACCGCC 508
DB 343 GCCGGAAGAGGCTGCGGGAAGAGAGCGCGCGCTCTTCAAGAGATGTCTCGCGG 402
QY 509 TGGCGCACTGTACACAGCAGGTCTGTCTGCTGCTGATCTCAAGCTGTGTCTGCTTGTCT 568
DB 403 TGGCCCACTGCGCCAGTCAAGCATGCTGTCTGCGGAGACCTGAAGCTTAGAGAGTGTCT 462
QY 569 TGGCTGACCGTGAAGAGAGAGAGTGTGTGTGAGAACTGAGAGACTCTGCTGTGTGTA 628
DB 463 TCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522
QY 629 CTGAGCGCAGATGATTCCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
DB 523 AGGGGGAAGATGATGTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582
QY 689 TACTCAGCTCACGCGGCTCTATCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 748
DB 583 TCTTCAACACACACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 642
QY 749 CGCTCTTCAACATGCTGCGCGGCGGCACTACCCCTTCCAGAGAGAGAGAGAGAGAGAGAG 808
DB 643 TGTCTTCAACACCTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
QY 809 TGGCAAGATTCGCGCGCGGCGGCTTACGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 868
DB 703 TCTCCAAATTTGGGCGTGGAGAGTTCGATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 762
QY 869 GTCTGTGCTGCTGCTCTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928
DB 763 GCTTCAATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 822
QY 929 TCTTGCACCCCTGAGCT 944
DB 823 TACTGCACCCCTGAGTT 938

```

RESULT 8

US-10-119-926-73

Sequence 73, Application US/10119926
 Publication No. US20030104413A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chenghua
 APPLICANT: Adundi, Vinod
 APPLICANT: Zhang, Jie
 APPLICANT: Ren, Feiyan
 APPLICANT: Zhao, Qing A.
 APPLICANT: Chen, Rui-hong
 APPLICANT: Wehrman, Tom
 APPLICANT: Zhou, Ping

APPLICANT: Xu, Chongjun
 APPLICANT: Yang, Yonghong
 APPLICANT: Xue, Aidong J.
 APPLICANT: Drmanac, Radoje T.
 TITLE OF INVENTION: No. US20030104413A1 Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 789C12B3CON
 CURRENT APPLICATION NUMBER: US/10/119,926
 CURRENT FILING DATE: 2002-04-09
 PRIOR APPLICATION NUMBER: 09/574,454
 PRIOR FILING DATE: 2000-05-19
 PRIOR APPLICATION NUMBER: 09/519,705
 PRIOR FILING DATE: 2000-03-07
 NUMBER OF SEQ ID NOS: 108
 SOFTWARE: pc_genes Version 1.0
 SEQ ID NO 73
 LENGTH: 1909
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (135) .. (944)
 US-10-119-926-73

Query Match 24.5%; Score 263.2; DB 15; Length 1909;
 Best Local Similarity 60.7%; Pred. No. 1,2e-59;

Matches 447; Conservative 0; Mismatches 288; Indels 1; Gaps 1;

```

QY 210 CTATGCTCTCTGAGAGCCGAGAGAGGCGGCGGCTACCGGCGCCCTGACCTGACCCCTAC 269
DB 103 CTACCTGCTGCTGCGCCCTAGCCGAGCGAGCATGTGTCCCGGCGCTGTGACATCCACAC 162
QY 270 AGGCACTGAGTATAC-CTGCAGAGTGTACCCGCTCCAGAGAGCCCTGCGCTGCTGAGAC 328
DB 163 TGGAGCGCGGAGTGGCTGCAAGGTCTTCCATTAAACATACAGAGCAAAATCAGGC 222
QY 329 CCTACGCGGCGGCTGCGCCGACAGCATGTGCTGCGCCCACTGAGTCTGCTGCTGTA 388
DB 223 CTTACATCCAGCTGCGCATGCGACAGCAATTACTGCGCATTTGTGAGAGTGTCTTGGGG 282
QY 389 CCCAGCTCTTACGCGCTTTTTCATCTCGACCAATGGGACATGCAACGCTGTGTGCGAA 448
DB 283 AAACCAAGGCTATGTCTTCTTGTAGAGAGACTTTGGGACATGCACTCTATGTGCGAA 342
QY 449 GCCGCAACCGTATCCCTGAGCGCTGAGCGCTGCTCTTCCGCGAGATGCGACCGCC 508
DB 343 GCCGGAAGAGGCTGCGGGAAGAGAGCGCGCGCTCTTCAAGAGAGATGTCTCGCGG 402
QY 509 TGGCGCACTGTACACAGCAGGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 568
DB 403 TGGCCCACTGCGCCAGTCAAGCATGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462
QY 463 TGTCTTCAACACCTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522
QY 569 TGGCAAGATTCGCGCGCGGCGGCTTACGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 628
DB 463 TCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582
QY 629 CTGAGCGCAGATGATTCCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
DB 523 AGGGGGAAGATGATCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582
QY 689 TACTCAGCTCACGCGGCTCTATCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 748
DB 583 TCTTCAACACACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 642
QY 749 CGCTCTTCAACATGCTGCGCGGCGGCACTACCCCTTCCAGAGAGAGAGAGAGAGAGAGAGAG 808
DB 643 TGTCTTCAACACCTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
QY 809 TGGCAAGATTCGCGCGCGGCGGCTTACGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 868
DB 703 TCTCCAAATTTGGGCGTGGAGAGTTCGATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 762
QY 869 GTCTGTGCTGCTGCTCTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928

```

```

Db      763 GCCTCATTCGACGCTCTTAAAGAGGAGGAGGCTCCGAGAGATCACTGCCCCGAGATCC 822
Qy      929 TCCTGCAACCCCTGGCT 944
Db      823 TACTGACACCCCTGCTT 838

```

RESULT 9

```

US-10-084-817-300
; Sequence 300, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon B. Pion
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 300
; LENGTH: 4336
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 234427.4
US-10-084-817-300

```

```

Query Match      23.6%; Score 253; DB 15; Length 4336;
Best Local Similarity 55.7%; Pred. No. 6e-57;
Matches 484; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

```

```

Qy      147 CCTGAGCCCACTACTGCTCCAGATGCTGAACCTGCTGAGCCAGCTCCCTCTTGG 206
Db      1459 CTTGGCTCTCCCGAGCCCGCCGAGACTCCGAACTTGTCGATGGCTTTCTGTATCGG 1518
Qy      207 GCCCTATGTCCTCTGAGAGCCGAGAGAGGCGGCGGCGCTACCGGCGCTGACAGCCG 266
Db      1519 GAAATACTTATTTGGAAGCTCTGAGAGGAGACACGTTTTCGTGCGCTGATCTGCA 1578
Qy      267 TACAGGCACTGATATACCTGCAAGGTGTAACCCGTCAGAAAGCCCTGCGCTGGA 326
Db      1579 CAGCGAGAGAGAGCTGCTGTCGAAAGTGTGATATCAGCTCTACAGGAATCCCTGGC 1638
Qy      327 GCCCTAAGCGGCGCTGCTCCCGGCAAGCAATGCGCTCGGCGCACTGAGTCTCTGCTGG 386
Db      1639 ACCGTGCTTTTGTCTGTCTGCTCATAGTAACTCAACAAATCACTGAAATTTATCTGG 1698
Qy      387 TACCAAGCTCTCTACGCTTTTTCATCTCGAGACCAATGGGAGATGACACAGCTGTGCG 446
Db      1699 TGAAGACCAAGCTATGTTGTTCTTTGAGGAAAGCTATGGGAGATGCTCTTCCTCG 1758
Qy      447 AAGCGCCACCGTATCCCTGAGCTGAGCTGCTGCTTTCCGCAATGCGCACCG 506
Db      1759 CACTGCAAGAAAGCTGAGAGAGAGAGGAGGAGCAACTGTTCTACAGATTCCTCGGC 1818
Qy      507 CCTGAGGCACTGTCACAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566
Db      1819 AGTGGCCCACTGTCATGAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1878
Qy      567 CTTGCTGACCGTGAAGAGAAAGCTGTGCTGAGAACTGAGAGACTCTGCTGCTGCT 626
Db      1879 CTTTAAAGACGAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1938
Qy      627 GACTGAGGCAAGATGATTCCTCTGAGGACAAAGCAGCTGCTGCTGCTGCTGCTGCTG 686
Db      1939 GCGGGAAGATGATGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1998

```

```

Qy      687 GATATCAAGTCAAGGAGGCTCTATCTGAGGCAAGGACGCGGATGCTGAGGCTTGGGCT 746
Db      1999 GATCTTGAACACGAGTGGACGCTACTCGGGCAAGAGCCGAGCTGTGAGGCTTGGAGGT 2058
Qy      747 GGGCTCTTACCATGCTGAGCGGAGCACTACCCCTTCCAGAGCTGAGAGCTGCTGCT 806
Db      2059 GATGCTGTACCATGTTGTGTGGGCGGTACCTTTTCATGACATTTGAACCCAGCTCCT 2118
Qy      807 CTTGCAAGATCCGCGCGGAGGCTACAGCTTGTGCTGCAAGGCTCTGAGCCCTGCTGCG 866
Db      2119 CTTGCAAGATCCGCGCGGAGGCTACAGCTTGTGCTGCAAGGCTCTGAGCCCTGCTGCG 2178
Qy      867 CTGTCTGTTGCTGCTCTCTTGTGAGGAGCAAGCTGAACGCTTACAGCCACAGGCT 926
Db      2179 GTGCTCATTCGGAAGATCTGTGCTGCGGAGCCCTCAGAGGCTGACTTCCGAGGAAAT 2238
Qy      927 CTTCTGCAACCCCTGCTGCGAGAGAGCCGATGAGCTTACGCTTACGCTTACGCTTAC 986
Db      2239 TCTGACATCTCTGCTTCTTCTACAGATTTTACGCTTCTGATTCAGATATGCTGCTAA 2298
Qy      987 CTGGAGAGCTGCCAGAGTGTCCCTGATG 1015
Db      2299 GGAAGTGTGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2327

```

RESULT 10

```

US-10-044-090-269
; Sequence 269, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 269
; LENGTH: 3280
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1820904CBI
US-10-044-090-269

```

```

Query Match      23.4%; Score 251.4; DB 14; Length 3280;
Best Local Similarity 55.6%; Pred. No. 1.6e-56;
Matches 483; Conservative 0; Mismatches 386; Indels 0; Gaps 0;

```

```

Qy      147 CCTGAGCCCACTACTGCTCCAGATGCTGAACCTGCTGAGCCAGCTCCCTCTTGG 206
Db      403 CTTGGCTCTCCCGAGCCCGCCGAGACTCGAACTTGTCGATTTGCTGTATCGG 462
Qy      207 GCCCTATGTCCTCTGAGAGCCGAGAGAGGCGGCGGCTTACCGGCGCTGACAGCTGCG 266
Db      463 GAAATACTTATTTGGAAGCTCTGAGAGGAGACACGTTTTCGTGCGCTGATCTGCA 522
Qy      267 TACAGGCACTGATATACCTGCAAGGTGTAACCCGTCAGAAAGCCCTGCGCTGCTGGA 326
Db      523 CAGCGAGAGAGAGCTGCTGTCGAAAGTGTGATATCAGCTGCTACAGGAATCCCTGGC 582
Qy      327 GCCCTAAGCGGCGCTGCTCCCGGCAAGCAATGCTGCTGCGGCACTGAGTCTCTGCTG 386
Db      583 ACCGTGCTTTTGTCTGCTCATAGTAACTCAACAAATCACTGAAATTTATCTCGG 642
Qy      387 TACCAAGCTCTCTACGCTTTTTCATCTGAGACCAATGGGAGATGACACAGCTGTGCG 446
Db      643 TGAAGCAAAAGCTTATGTTGTTCTTTGAGGAAAGTATGGGACATGCAATTCCTTCCTCG 702
Qy      447 AAGCGCCACCGTATCCCTGAGCTGAGGCTGCTGCTTCCGCGCAATGAGCCACCG 506

```

Db	703	CACCTGGCAAGAAAGCTGAGAGAGAGAGAGAGGACAGACTGTTCTTACAGATTTGGCTTCGGC	762
Oy	507	CCTGGCGCACTGTATCACAAGCAACGGTCTGGTCTCTGGTGAATCTCAAGCTGTGTCCGTTGT	566
Db	763	AGTGGCCCACTGCCTATGACGGGGGGCTGGTGTCTGGCGGACCTCAAGCTGTGGAAATTCAAT	822
Oy	567	CTTGGCTACCGCTGAGAGAGAAAGAACTGGTGTCTGGAGAACTTGAAGAGACTCCGTGGTGT	626
Db	823	CTTTAAGAACGAAGAGAGAACTCTGGGCTCAAGCTGGAAAGCTTGAAGAGCGCTTACATTCT	882
Oy	627	GACTGGGCGAGATGATATTCCTCTGTGGGACAAAGCACCGCTGCCAGCTTACGTGGGACCTGA	686
Db	883	GCGGGGAATATGATATTCCTCTCTCGAGCAAGATATGGCTGCCGGCTTAAGTAAGCCAGA	942
Oy	687	GATATCAAGCTCACGGGCTCATTTCTGGGCAAGGACCGGATGTCTGGAGCTTGGGCGT	746
Db	943	GATCTTGAACACACAGTGGCAGCTTACTCTGGGCAAAAGCACCGGACGTGTGGAGCTGGGGGT	1002
Oy	747	GGCGCTCTTACCAATGCTGGCCGGCCACTAACCCCTTCCAGACCTGGAGCTGTCTGTCT	806
Db	1003	GATGCTGTACACCAAGTGGTGGGGCGGATACCTTTTCCATGACATTTGAACCCAGCTCCCT	1062
Oy	807	CTTGGCGCAAGATTCGCGCGCGGGGCGCTACGCGCTTTGGCTGACGAGCGTCTGGGCCCTGGCCG	866
Db	1063	CTTACAGCAAGACCCGGCGCTGGCGCAATTCACATTTCCAGAGACTTGTGTGGCCCAAGCCAA	1122
Oy	867	CTGTCTGGTTGCGTGCCTCTTTCGTGGGAGGCGACGTGAACGGCTCAACAGCCACAGGAT	926
Db	1123	GTGCTTCATCCGAAGCATTTGTGCGTCCGGAGGCCCTCAGAGCGGCTGACCTGTGCAGGAAT	1182
Oy	927	CCTCCTGACCCCTGTGGCTGCGACAGAGACCCGATGCCCTTAAGCCCAACCCGATCCCATCT	986
Db	1183	TCTGACACCATCTTGGTGTTCATACAGATTTTATGCGCTTCGAATTCAGACATATGTGTAA	1242
Oy	987	CTGGGAGCGTCCCAAGGTGTCCCGATG	1045
Db	1243	GGAAGTGTACACAGCTGTGTGCCGAGCG	1271

```

RESULT 11
US-10-240-965-168
Sequence 168, Application US/10240965
Publication No. US20030165924A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: SHIFFMAN, Dov
APPLICANT: SOMOGYI, Roland
APPLICANT: LAMN, Richard M.
APPLICANT: SEILHAMER, Jeffrey J.
APPLICANT: PORTER, Gordon J.
APPLICANT: MIKITA, Thomas
APPLICANT: TAI, Julie
TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
FILE REFERENCE: PA-0025 PCT
CURRENT APPLICATION NUMBER: US/10/240,965
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: 60/195,106
PRIOR FILING DATE: 2000-04-05
NUMBER OF SEQ ID NOS: 276
SOFTWARE: PERL Program
SEQ ID NO 168
LENGTH: 3324
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030165924A1 197301.4
FEATURES:
NAME/KEY: unsure
LOCATION: 254..378
OTHER INFORMATION: a, c, g, or other
US-10-240-965-168

```

Query Match	23.0%	Score 247	DB 13	Length 3324
Best Local Similarity	61.1%	Pred. No. 2,38-55		
Matches	400	Conservative	0	Mismatches 255
			Indels	0
			Gaps	0
QY	290	AGGTGTACCCCGTTCAGAGAACCCCTGGCCGCTGTGAGCCCTTACGCGCGCTGCCCGCC	349	
DB	635	AGGGTGTTCCTATTAAACATCACAGACAAATAACAGCCCTTATCATCAGCTGGCCATGCG	694	
QY	350	ACAAGCATGTGTGCTCGGCGCCCATGAGGTCTCTGGCTGTGATACCCAGCTCTCTACGCTTTT	409	
DB	695	ACAGCAACATTACTTGCCATTGTGTGAAATGATCTTTGGGAAACCAAGGCTTATGTCTTCT	754	
QY	410	TCACTCGGACCCCATGTGGGACATGACAGCCCTGTGTGCAGAACCGCACCCGTATCCCTGAGC	469	
DB	755	TTTGAAAGGACTTTTGGGGACATGCACTCTATGTGTGCAAGCCGGAAGAGCGCTGCGGGAAG	814	
QY	470	CTGAGGCTGCGCTGTCTCTTCCGACAGATGTGCACCGGCTTGGGCCCATGTATACACAGACG	529	
DB	815	AGGAAGCGCGCGGCTCTTCAAGAGATGTCTCCGCGTGCACATGCGCACCACTGATG	874	
QY	530	GTCGTGTCTGTGATCTCAAGCTGTGTGCTTGTCTTGTGCTGACCGTGTAGAGAGA	589	
DB	875	CCATCGTGTGTGGGACCTGAAAGCTTATGAAAGTTGATCTTCTCACGAGAGAGAAACC	934	
QY	590	AGCTGTGTGTGAGAACCTGTGAGGATCTCTGCTGTGACTGTGGCCAGATATTCCTGT	649	
DB	935	AGCTTAACTAGAAAGTCTAGAAACACACACATATTAAGAGGGGAAATGATCTTTGT	994	
QY	650	GGGACAAAGCAGCGCTGCCACGCTACCTGTGGACCTGTAGATACTGAGCTCACGGGCTCAT	709	
DB	995	CAGCAAAACATGTGCTGCCAGCTACGTTGAGCCCTTGATCTCAACACACCTGGAGCT	1054	
QY	710	ACTGTGGCAGAGCAGCCGATGTCTGAGAGCTGTGGCGCTGTGCGCTTTTCCATGCTGCGG	769	
DB	1055	ACTCCGGAAGGCTGCGGACGTTTGTGAGCCCTGGGGGTGATGTCTTACACCCCTTGTGTG	1114	
QY	770	GCACATAACCCCTTCAAGACCTGGAGGCTGTGCTGTCTTTCGAGCAAGATTCGCGCGGGGCT	829	
DB	1115	GACGATACCCCTTCCATGATCTCAGAACCCCAATGCTCTTTTCTCMAAATTCGGGTGTGAC	1174	
QY	830	CTTACGCTTGTGCTGTGAGGCTCTCGGCGCCCTGTGCGCTGTCTGTGCTCTCTTTC	889	
DB	1175	AGTTCCTCATTTCTGACACATTTCCCAAGCAGATGTCTTATTCGACGCTCTTTGA	1234	
QY	890	GTCGAGGACGAGCTGAACGGGTCAACAGCCACAGGCTCTCTGACACCCCGGCT	944	
DB	1235	GACGGAGGCTCTCGAAGACTCATCTGCCCCGAGATCTTACTAGCAACCCCTGTGTT	1289	

```

RESULT 12
US-10-228-263-1
? Sequence 1, Application US/10228263
? Publication No. US2003009985A1
? GENERAL INFORMATION:
? APPLICANT: li, jing
? TITLE OF INVENTION: AMPLIFIED GENE INVOLVED IN CANCER
? FILE REFERENCE: 38002-0034
? CURRENT APPLICATION NUMBER: US/10/228, 263
? PRIOR FILING DATE: 2002-12-04
? PRIOR APPLICATION NUMBER: US 60/330, 797
? PRIOR FILING DATE: 2001-10-31
? PRIOR APPLICATION NUMBER: US 60/314, 655
? PRIOR FILING DATE: 2001-08-27
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 1
? LENGTH: 818
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (788)..(788)
? OTHER INFORMATION: N can be C or T

```

FEATURE:
NAME/KEY: CDS
LOCATION: (56)..(676)
OTHER INFORMATION:
US-10-228-263-1

Query Match 21.6%; Score 232.4; DB 15; Length 818;
Best Local Similarity 64.1%; Pred. No. 1, 5e-51;
Matches 350; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

```
Oy 399 CTACGCCCTTTTCTACTCGGACCCATGAGGACATGACACAGCCTGTGTGCGAAGCCGACCG 458
Db 25 CTATGCTTTCTTTGAGAAAGACTTTGGGACATGACCTCTATGTGCGAAGCCGGAAG 84
Oy 459 TATCCTGAGCCTGAGGCTGCTGCTCTTCCGCCAGATGCGACCGCCCTGGCCGACTG 518
Db 85 GCTGGGAGAAAGAGAGCCGCGCCGCTCTTCAAGACAGATTTCTCCGCGCTGCCACTG 144
Oy 519 TCACGACAGCGTGTGTCTGTCTGTATCTCAAGCTGTGTGCTTTGTCTTGTCTGACCG 578
Db 145 CACACAGTCAAGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 204
Oy 579 TGAGAGAAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 638
Db 205 GAGAGAAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 264
Oy 639 TGATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 698
Db 265 TGATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 324
Oy 699 ACGGAGCTCATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 758
Db 325 CACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 384
Oy 759 CATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 818
Db 385 CTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 444
Oy 819 CCGCGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 878
Db 445 TCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 504
Oy 879 CTGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 938
Db 505 CAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 564
Oy 939 CTGCT 944
Db 565 CTGCT 570
```

RESULT 13
US-10-228-263-3
Sequence 3, Application US/10228263
Publication No. US200309985A1
GENERAL INFORMATION:
APPLICANT: LI, JING
TITLE OF INVENTION: AMPLIFIED GENE INVOLVED IN CANCER
FILE REFERENCE: 38002-0034
CURRENT APPLICATION NUMBER: US/10/228,263
CURRENT FILING DATE: 2002-12-04
PRIOR APPLICATION NUMBER: US 60/330,797
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: US 60/314,655
PRIOR FILING DATE: 2001-08-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 621
TYPE: DNA
ORGANISM: Homo sapiens
US-10-228-263-3

Query Match 20.7%; Score 222.2; DB 15; Length 621;
Best Local Similarity 64.5%; Pred. No. 7, 2e-49;
Matches 332; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

```
Oy 430 ATGACAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 489
Db 1 ATGACAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
Oy 490 CGCCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 549
Db 61 AAGCAGATTTGTCTCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Oy 550 AAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
Db 121 AAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Oy 610 GAGACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 669
Db 181 GAGACACACATATATAGAGGAGATATATATATATATATATATATATATATATATAT 240
Oy 670 GCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 729
Db 241 GCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Oy 730 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 789
Db 301 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Oy 790 TCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 849
Db 361 TCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Oy 850 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 909
Db 421 ATTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Oy 910 CTGACAGCAGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 944
Db 481 CTGACAGCAGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 515
```

RESULT 14
US-09-925-301-525
Sequence 525, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: ROSEN ET AL.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 525
LENGTH: 562
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (515)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (526)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (557)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-525

Query Match 20.6%; Score 220.8; DB 9; Length 562;
 Best Local Similarity 96.1%; Pred. No. 1.7e-48;
 Matches 219; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGCGAGCCACCCCTCTGCTGCTCTGCGGGTTCCCTGTCAGAAAGCGGTTGAG 60
 DB 322 ATGCGAGCCACCCCTCTGCTGCTCTGCGGGTTCCCTGTCAGAAAGCGGTTGAG 381
 QY 61 TTGATGACAACTTAGATACCGAGCGTCCGTCAGAAAAGAGCTGAAAGTGGCCCG 120
 DB 382 TTGATGACAACTTAGATACCGAGCGTCCGTCAGAAAAGAGCTGAAAGTGGCCCG 441
 QY 121 CCCAAGTCCCGCCCTGCTGTTGGCCCTGAGGCCCACTACTGCTCAATGTGAACT 180
 DB 442 CCCAAGTCCCGCCCTGCTGTTGGCCCTGAGGCCCACTACTGCTCAATGTGAACT 501
 QY 161 GCTGTGCGCACTGCGCTCCGCTCTTGGGCGCTATAGTCTCTGAGGCC 228
 DB 502 GCTGTGCGCACTGCGCTCCGCTCTTGGGCGCTATAGTCTCTGAGGCC 549

RESULT 15

US-10-291-172-492
 ; Sequence 492, Application US/10291172
 ; Publication No. US20030228584A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HySeq, Inc
 ; TITLE OF INVENTION: No. US20030228584A1 Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 21272-045
 ; CURRENT APPLICATION NUMBER: US/10/291,172
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 09/693,267
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 09/665,363
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 09/616,847
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 09/596,193
 ; PRIOR FILING DATE: 2000-06-17
 ; PRIOR APPLICATION NUMBER: 09/574,454
 ; PRIOR FILING DATE: 2000-05-19
 ; PRIOR APPLICATION NUMBER: 09/519,705
 ; PRIOR FILING DATE: 2000-03-07
 ; NUMBER OF SEQ ID NOS: 752
 ; SEQ ID NO 492
 ; LENGTH: 1943
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-291-172-492

Query Match 19.3%; Score 207.6; DB 12; Length 1943;
 Best Local Similarity 60.2%; Pred. No. 5.2e-45;
 Matches 396; Conservative 0; Mismatches 259; Indels 3; Gaps 3;

QY 290 AGGTGTACCCCGTCCAGGAAGCCCTGCGCGCTGAGCCCTACGCGCGCTGCCCCGC 349
 DB 221 AGGTGTTCCTTAACTAACTACAGAGCAAAATCAGGCTTACATCCAGCTGCATGCG 280
 QY 350 ACAAGCATGTGCTCGGCCCACTGAGTCTCTGCTGATACCGAGCTCTTACGCTTTT 409
 DB 281 ACAGCAATTACTGAGCATGTGAGAGTATCTTGGGGAACCAAGGCTATATGCTTCT 340
 QY 410 TCACCTCGAGCCATAGGAGCATGCAAGCTGTGCGAAGCCGCCACGCTATCCCTGAGC 469
 DB 341 TTGAGAAAGACTTTGGGAGCATGCACTCTATGTGCAAGCCGGAAGAGGCTGCGGAAG 400
 QY 470 CTGAGGCTGCGGTGCTCTTCCGCGAGATGAGCCAGCCCTGCGGAGCTGTGACAGCAG 529
 DB 401 AGGAAGCCGCGCGCTTTCAAGCAGATGTCTCCGCGCTGCGCACTGCAAGTCAAG 460
 QY 530 GTCTGCTCTGCTGATCTCAAGCTGTGCTGCTTTGCTTGTGCTGACCGTGAAGAAAG 589
 DB 461 CCATGCTGCTGCGGAGCCTGAAGCTTAAGAAAGTGTCTTCTTCCAGGAGAGAAAGCC 520

QY 590 AGCTGTGCTGAGAACTGTGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
 DB 521 AGCTTAGACTAGAAAGTCTAGAAAGACACACATATATGAAGGAGGAAATGATGCTTGT 580
 QY 650 GGGACAAAGACGCGGCCAGCTACGTGAGACCT- GAGATATCTAGCTCAGGGGCTCA 708
 DB 581 CAGACAAACATAGCTGCCCCAGCTTACGTAGAGCCCTGAGATCTTCAACACACTGGGACC 640
 QY 709 TACTGCGCAAGAGCAGCCGATGTCTGA- GCTGAGCGTGGGCTCTTCAACATGCTGG- 766
 DB 641 TACTCGAAAGGCTGCGGACGTTTGAAGTGGGGGGGGTGAATGCTTACACCTTGTGT 760
 QY 767 CCGGCCACTACCCCTTCCAGAGACTGGAAGCTGCTCTGCTTGGCAAGATCGCGCG 826
 DB 701 TGGAGAGATACCCCTTCCATGACTCAGACCCAGTGCCTTTTCTCAAAATTCGCGGTG 760
 QY 827 GGGCTTACGCTTGTGCTGAGGCTCTGAGCCCTGCGCCCTGCTGCTGCTGCTGCTC 886
 DB 761 GACAGTTCGCAATTCCTGAGCAATTTCCCAAGCCAGTGCCTGATTCGACGCTCT 820
 QY 887 TTGTCGAGGAGCCAGCTGAACGGCTCAACAGCCAGGCAATCTCTGACACCCCTGCT 944
 DB 821 TGAAGAGGAGGCTCTCCAGAGACTCATGCCCCCGAGATCTTACTGCAACCCCTGCT 878

Search completed: January 16, 2004, 20:19:10
 Job time : 373.799 secs

Db	488	GCTCGGCCCATGAGAGTCTGGCTGTATCCAGGCTCTCTAAGCTTTTCACTGGAGC	547
Oy	421	CATGGGGACATGACACGCTGTGTGGAAAGCCGCACCGTATTCCTGAAGCTGAAGCTGCC	480
Db	548	CATGGGGACATGACACGCTGTGTGGAAAGCCGCACCGTATTCCTGAAGCTGAAGCTGCC	607
Oy	481	GTGCTCTTCCGGCAGATGAGCCACCGCCCTGGGGGCACTGTACACAGCAAGGTCTGGTCC	540
Db	608	GTGCTCTTCCGGCAGATGAGCCACCGCCCTGGGGGCACTGTACACAGCAAGGTCTGGTCC	667
Oy	541	CGTATCTCAAGCTGTGTGCTTTGTCTTCGCTGAACCGTAGAGGAAGAGCTGTGCTG	600
Db	668	CGTATCTCAAGCTGTGTGCTTTGTCTTCGCTGAACCGTAGAGGAAGAGCTGTGCTG	727
Oy	601	GAGAACCTGGAGGACTCCGTCGCTGTGACTGGGCCAGATGATTCCTGTGGGACAAGAC	660
Db	728	GAGAACCTGGAGGACTCCGTCGCTGTGACTGGGCCAGATGATTCCTGTGGGACAAGAC	787
Oy	661	GCGTGCCAGGCTTACGTGGGACCTTGAGATACTCAAGCTCACGGGCTCTACTCGGGCAAG	720
Db	788	GCGTGCCAGGCTTACGTGGGACCTTGAGATACTCAAGCTCACGGGCTCTACTCGGGCAAG	847
Oy	721	GCAGCGCATGTCTGAGAGCTTGAGGCTGGGCGCTTACACATGCTTGCCGGGCCATACCC	780
Db	848	GCAGCGCATGTCTGAGAGCTTGAGGCTGGGCGCTTACACATGCTTGCCGGGCCATACCC	907
Oy	781	TTCCAGAGCTGGAGCTGTCTGTCTTTCGGCAGAGATCCGCGGGGGCTTACGCTTGG	840
Db	908	TTCCAGAGCTGGAGCTGTCTGTCTTTCGGCAGAGATCCGCGGGGGCTTACGCTTGG	967
Oy	841	CTTGAGGCTTCTCGGAGCCCTTGCCCGGCTGTCTGTTGCTGTGCTCTTCTGTGGAGGCCA	900
Db	968	CTTGAGGCTTCTCGGAGCCCTTGCCCGGCTGTCTGTTGCTGTGCTCTTCTGTGGAGGCCA	1027
Oy	901	GCTGAAGGGCTCACAGGCCACAGGATTCCTCCGGAACCCCTGGCTGGACAGGACCCGA	958
Db	1028	GCTGAAGGGCTCACAGGCCACAGGATTCCTCCGGAACCCCTGGCTGGACAGGACCCGA	1085

```

RESULT 2
US-09-220-132-135/c
! Sequence 135: Application US/09220132
! Patient No. 6506607
! GENERAL INFORMATION:
! APPLICANT: Shyjan, Andrew W.
! TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
! TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
! FILE REFERENCE: 07334-074001
! CURRENT APPLICATION NUMBER: US/09/220,132
! CURRENT FILING DATE: 1998-12-23
! PRIOR APPLICATION NUMBER: US 60/079,303
! PRIOR FILING DATE: 1998-03-25
! PRIOR APPLICATION NUMBER: US 60/066,821
! PRIOR FILING DATE: 1997-12-24
! NUMBER OF SEQ ID NOS: 191
! SOFTWARE: FastSeq for Windows Version 4.0
! SEQ ID NO 135
! LENGTH: 2559
! TYPE: DNA
! ORGANISM: Homo sapiens
US-09-220-132-135

```

Query Match Similarity 39.7%; Score 426.2; DB 4; Length 2559;
Best Local Similarity 93.8%; Pred. No. 1.7e-87;
Matches 487; Conservative 0; Mismatches 28; Indels 4; Gaps 4;

Db	1514	TCCTGCGGTGCTGATCGGAGCCAGATGATTCCCTGTG-6GA	CAAGCAGCGGTGCGCAGCTCAC	1516		
Oy	676	GTGGAGCCTTGAGATACTCAAGCTCA	CGGGGCTCATACTCGGGCAAGGACCGGATGTCTGG	735		
Db	1515	GTGGAGCCTTGAGATACTCAAGCTCA	CGGGGCTCATACTCGGGCAAGGACCGGATGTCTGG	1456		
Oy	736	AGCTTGGGCGTGGGCGTCTTCA	CCATAGCTGTGGCCCATACCCCTTCCAGGATTCGGAG	795		
Db	1455	AGCTTGGGCGTGGGCGTCTTCA	CCATAGCTGTGGCCCATACCCCTTCCAGGATTCGGAG	1396		
Oy	796	CCTTCTCTGCTCTTCCGCAAGATTC	CGCGCGGGGCGCTTAACGCTTGGCCCTGCAAGGCGTCTCG	855		
Db	1395	CCTTCTCTGCTCTTCCGCAAGATTC	CGCGCGGGGCGCTTAACGCTTGGCCCTGCAAGGCGTCTCG	1336		
Oy	856	GCCCCCTGCGCGTGTCTGATTCTG	GTGCGCTCTTCTGTGGGGAGCCAGCGTGAA	CGGCTCAC	915	
Db	1335	CCCCCTGCGCGTGTGTGATTCTG	GTGCGCTCTTCTGTGGGGAGCCAGCGTGAA	CGGCTCAC	1276	
Oy	916	GCACAAGGCAATCCTCTGCAACCC	CTGTGCTGCGAGACCGGATGACCCTTA	AGCCCCAAC	975	
Db	1275	G-CA	CGAGATCTCTCTGCAACCCCTGTGCGAGACA-G	ACCAGATGCCCTTA	AGCCCCAAC	1218
Oy	976	CGATCCCATCTCTTGGGAGGCTGC	CCCAAGTGTCCCTGATGGAATCTGGGCTTGGA	CGAAGCC	1035	
Db	1217	CGATCCCATCTCTTGGGAGGCTGC	CCCAAGTGTCCCTGATGGAATCTGGGCTTGGA	CGAAGCC	1159	
Oy	1036	AGGAGAGAGGAGGAGACAGAGAA	GTGTCTGTATGGC	1074		
Db	1158	AGGAGAGAGGAGGAGACAGAGAA	GTGTCTGTATGGC	1120		

RESULT 3
 US-08-146-421-4/c
 : Sequence 4, Application US/08146421
 : Patent No. 5543489
 :
 : GENERAL INFORMATION:
 : APPLICANT: BREWER, GARY
 : TITLE OF INVENTION: DNA SEQUENCE ENCODING A POLYPEPTIDE WITH
 : TITLE OF INVENTION: ANTI-TUMOR PROPERTIES
 : NUMBER OF SEQUENCES: 9
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: DILMORTH & BARRESE
 : STREET: 4350 LA JOLLA VILLAGE DRIVE, SUITE 300
 : CITY: SAN DIEGO
 : STATE: CALIFORNIA
 : COUNTRY: U.S.A.
 : ZIP: 92122
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/146,421
 : FILING DATE: 29-OCT-1993
 :
 : CLASSIFICATION: 514
 : ATTORNEY/AGENT INFORMATION:
 : NAME: PEPPER PH.D., FREDRICK W.
 : REGISTRATION NUMBER: 31,286
 : REFERENCE/DOCKET NUMBER: 489-2
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 619-546-4410
 : TELEFAX: 619-453-2839
 : INFORMATION FOR SEQ ID NO: 4:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2562 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 246..1106

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1146:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G1480860
US-09-016-434-1146

Query Match 7.9%; Score 85; DB 4; Length 1302;

Best Local Similarity 48.1%; Pred. No. 1.5e-10;

Matches 241; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

455 ACCGATATCCCTGAGCTGAGCTGCGCTCTTCGCGAATGAGCAACCGCCCTGGGCGC 514
437 AGCGTTTCCCATGTGCGACGAGCCGCGGTACTTCTGTACGCTGATTGACGCGCTGGAGT 496
515 ACTGTACACGACGAGCTGTGCTGTGCTGATCTCAAGCTGTGTGCTTTGCTTGGCTG 574
497 ACCTGCATAGCAGGAGCATTTGTGCAACAAGACATCAAGCGGGAACCTGTCTCACCAC 556
575 ACCGTGAGAGGAAGAGCTGTGTGTGTGAGAACTGTGAGACTCTCGCTGTGACTGGGC 634
557 CCGGTGGGACACCTCAAAATCTCCGACTGTGGGTGGCGGAGGCACTGCAACCGTTGCGCG 616
635 CAGATGATTCCTCTGTGGGCAAGACGCGTGGCCAGACCTAAGTGGGACTGTGATATCTCA 694
617 CGAGACGACCTGCGGACGACGAGGCTCCCGGCTTTCAGCGCCCGAGATTGGCA 676
695 GCTCACGGGCTCATATCTCGGCAAGGACGAGCTGTGAGAGCTGTGGCGGTGGCGCTCT 754
677 ACGGCTGTGACACTTCTCCGCTTCAAGTGTGAGATCTGTGCGGTGGGAGTCACTCTCT 736
755 TCACATGCTGAGCGGCACTACCTCTTTCAGAGACTCGAGAGCTCTCTCTCTTTCGCA 814
737 ACAATATACCAACGAGGTGTGATCCCTTGAAGGGGACAACATCTACAAGTTGTTGAGA 796
815 AGATCCGCGCGGAGCTACGCTTGTGCTGAGGCTCTCGGCGCTGCGCGCTGTGAG 874
797 ACATGGAAGGAGAGCTACGCACTCCGGGAGACTGTGGCCCTCCGCTCTTGTGACCTTC 856
875 TTCGCTGCTCTTCTGTGAGGAGCACTGAACGAGCTCAAGGCAACAGGACTCTCTTCG 934
857 TGAAGGAGATCTTGTGATGACGAACGAGGCAAGAGGTTCTCATTCGCGAGATCCGCGAGC 916
935 ACCCTGGCTGGAAGAGACC 955
917 ACAAGTGTTCGGAAGAAAC 937

RESULT 6

US-08-749-902-2
Sequence 2, Application US/08749902

Patent No. 5985635

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Goli, Surya K.

APPLICANT: Hillman, Jennifer L.

TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESS: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/749,902

FILING DATE: Filed Herewith

PRIOR APPLICATION NUMBER:

APPLICATION DATA:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0150 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1466 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: Consensus

US-08-749-902-2

Query Match 7.6%; Score 81.6; DB 2; Length 1466;

Best Local Similarity 47.3%; Pred. No. 9.1e-10;

Matches 237; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

455 ACCGATATCCCTGAGCTGAGCTGCGCTCTTCGCGAATGAGCAACCGCCCTGGGCGC 514
536 AGCGTTTCCCATGTGCGACGAGCCGCGGTACTTCTGTACGCTGATTGACGCGCTGGAGT 595
515 ACTGTACACGACGAGCTGTGCTGTGCTGATCTCAAGCTGTGTGCTTTGCTTGGCTG 574
536 ACCTGCATAGCAGGAGCATTTGTGCAACAAGACATCAAGCGGGAACCTGTCTCACCAC 655
575 ACCGTGAGAGGAAGAGCTGTGTGTGTGAGAACTGTGAGACTCTGCGTGTGACTGGGC 634
656 CCGGTGGGACACCTCAAAATCTCCGACTGTGGGAGTGGCCGAGCACTGACACCGCTTGGCG 715
635 CAGATGATTCCTCTGTGGGCAAGACGCGTGGCCAGCTTACGTGGAACCTGTGATATCTCA 694
716 CGAGACGACCTGCGGACGACGAGGCTCCCGGCTTTCAGCGCCCGGAGATTGGCA 775
695 GCTCACGGGCTCATATCTCGGCAAGGACGAGCTGTGAGAGCTGTGGCGGTGGCGCTCT 754
776 ACGGCTGTGACACTTCTCCGCTTCAAGTGTGAGATCTGTGCGGTGGGAGTCACTCTCT 835
755 TCACATGCTGAGCGGCACTACCTCTTTCAGAGACTCGAGAGCTGTGCTCTTTCGCA 814
836 ACAATATACCAACGAGGTGTGATCCCTTGAAGGGGACAACATCTACAAGTTGTTGAGA 895

QY 468 GCCTGAGCTGCTGCTCTTCCGCAATGCGCACCGCCCTGCGCAGCTGTCACGACA 527
DB 396 GCCAAGAGTGGCTACTACCTGGCGAGATCTCTTCTGGCTTCAGTACTGACCGAG 455
QY 528 CGGTCTGCTCTGCTGATCTCAAGCTGTGTGCTTGTCTTCTGCTGACCGGTGAGGAA 587
DB 456 CGGCACTTGGACCGGAGACTCAAGTTGGGAAATTTTTCATAC-----TGAGAACAT 509
QY 588 GAACTGTGTGCTGAGAACCTGAGAGACTCTGCTGCTGCTGAGCGGCAAGATATCCCT 647
DB 510 GGAATGAAAGTGTGGGAGATTTTGGGCTGGAGCGCGGTTGAGACCTCCGAGAGAGAA 569
QY 648 GTGGGACAAAGACCGCTGCCAGCTTACGTGGAGCTGAGATCTCAAGCTCACGGGCTTC 707
DB 570 GAAAGACATGTGTGGACCGCCCAATATGTGTGCTCCAGAAATGCTCTGAGACGGGCGCA 629
QY 708 ATACTGGGCAAGGCGGATGTCTGAGGCTGGGCGGTGGGCTCTTACCATATGCTGGC 767
DB 630 C-----GGCCTGAAAGCGGATGTATGTATCATCTGGGCTGTATGATACAGCTGCTG 683
QY 768 CGGCACTACCCCTTCCAGAGACTCGAGGCTGTCTGCTTCTGCGAAGATCCGCGCGG 827
DB 684 CGGAGCCCTCTCTTGAAGACGCTGACCTGAGAGAGAGTACCGCTGATCAAGAGAT 743
QY 828 GGCCTACGCTTCTGCTGACGAGCTCTGGGCGCTGCTGCTGCTGCTGCTGCTGCT 887
DB 804 TCGGGCTGACCCCGAGAGCGGCGCTTATTAGCAAGATCTGCGGCAATGCTTTTAC 863
QY 948 ACAGAGCCGAGTCCG 963
DB 864 CAAGGCTACACCCG 879

RESULT 9

US-09-579-664B-4
Sequence 4, Application US/09579664B
Patent No. 6514719
GENERAL INFORMATION:
APPLICANT: Immunex Corporation
APPLICANT: Bird, Timothy A.
APPLICANT: Virca, G. Duke
APPLICANT: Martin, Unja
APPLICANT: Anderson, Dirk M.
TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
FILE REFERENCE: 2923-A
CURRENT APPLICATION NUMBER: US/09/579, 664B
CURRENT FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 2902
TYPE: DNA
ORGANISM: Mus musculus
US-09-579-664B-4

Query Match 5.1%; Score 54.6; DB 4; Length 2902;
Best Local Similarity 49.7%; Pred. No. 0.0013;
Matches 231; Conservative 0; Mismatches 219; Indels 15; Gaps 3;
QY 451 CGCCACCTATCTCCGAGCTGAGCTGCTGCTCTGCGCCAGATGCGCACCGGCTG 510
DB 564 CGGCCACGCTGAGTGAAGGGGAGCGCCAGGCAATTTCTTCGACAGATGCTGTGCTG 623
QY 511 GCGCACTGTACACGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
DB 624 CACTACTGTCCACGAGAGCGGAGTGTGTACCGAGATCTCAAGCTGAGAAACATCTTCTA 683
QY 571 GCTGACCGTGAAGGAAAGTGTGTGTGAGAACTCTGAGAGACTCTGCGGTGACT 630

DB 684 GATGCCAATGAAACATCAAGATTGCTGACTTTGGCT-----CTCCAACTGTAC 734
QY 631 GGGCCAGATGATTCCTGTGGGACAAAGACGCGTGGCCAGCTTACGTGGAGCTGAGATA 690
DB 735 CACAAAGGAAATCTCTCCAGACGTTCTGTGGAGGCTCTCTACCGCTGCGCTGAGATA 794
QY 691 CTGAGCTCAAGGAGCTCATATCTCGGGCAAGGACCGGATGTCTGAGAGCTGGGCGG 750
DB 795 GTCAACGGGAAACCC---TATGTGGGCCCAAGAGTGAAGAGTGTCTGCTGCTGCTC 851
QY 751 CTCTTACGATCTGGCGCGGCACTACCCCTTCCAGAGCTCGGAGGCTGTCTGCTTTC 810
DB 852 CTGTACATCTGGTGACAGACATGCTCTTTGAGCGGAGATCATTAACAATGCTG 911
QY 811 GCGAAGATCGCGCGCGGCTTACGCTTGTGCTGAGGCTCTGCGGCGCTGCGCGCTGT 870
DB 912 AGCAAAATCAATGAAAGGGGCTTACCGTGAAGCGCC---CAAGCGTCCGATGCTGTGGC 968
QY 871 CTGCTTCTGCTGCTCTCTTCTGCTGCGGAGCGCAAGCTGAACGCTCA 915
DB 969 CTGATCGGCTGCTGTATGTGAACCCACCGCTGCGGCGCACA 1013

RESULT 10

US-08-125-468-1/C
Sequence 1, Application US/08125468
Patent No. 5589385
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Lotvin, Jason A.
APPLICANT: Strathy, Nancy
APPLICANT: Fancini, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmids
TITLE OF INVENTION: useful therein
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSER: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,468
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Isevedor, Bette J.
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31,255-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 831-3241
TELEFAX: (201) 831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-125-468-1

Query Match 4.8%; Score 51.2; DB 1; Length 30001;
Best Local Similarity 44.2%; Pred. No. 0.011;
Matches 212; Conservative 0; Mismatches 268; Indels 0; Gaps 0;
QY 390 CCAAGCTCTTAAGGCTTTTCACTGGAACCAATGAGGAGATGACAGCGCTGTGGAAG 449

Db	22561	CGAAGGCTCAACGGATGTTGGCTTGGCGGCTGTGGACAGAGCAAGAGGAAGTCTCT	22502
Qy	450	CCGCGACCGTATCCCTGAGCCCTGAGCGTGGCTCTTCCGCAGATAGCAACGGCCT	509
Db	22501	GCTGGTCCGGACCGGATGGGGCTCAAGCGCGCTCTACTACTGAGCCAGACGGCGGCGGT	22442
Qy	510	GGCGCACTGTCACAGACCGGTCTGGTCTGGCGTATCTCAAGCTGTGGCTTGTCTT	569
Db	22441	GCGCTTGGGCTCGAGGCCAAGGCGCGTCTGGCGAGACCGAACCTGTCTCCGCGGGTCCG	22382
Qy	570	CGCTGACCGTGAAGGAAGGCTGGGCTGGAGAACTGGAGGACTCTGCGGTGTGAC	629
Db	22381	CCCGGAGGGCGTGGCGAGGTGCTGACATGTAAGAACGCCCGAAGCGGCGGTCTTCTC	22322
Qy	630	TGGCGCAGATGATTTCCCTGTGGGACAAAGCACGCGTCCGACCTTACGTGGACCTGAGAT	689
Db	22321	CGGGCTGTGGAGGTCCGTCCGGGTCAAGCTGTGCGGGTGGGCGGGGGGTGGCGCG	22262
Qy	690	ACTCAGCTCAACGGGCTCATATCTGGGAGGAAGGACCGATGTCTGGAGCGTGGCGTGGC	749
Db	22261	GCAGACGTACTGGGGGCTGGAGGCCCGCGAGCACACCGACGACTCTCCGACACATCGG	22202
Qy	750	GCTTTTCAACATGTGGCCGGCCACTACCCCTTCAGAGACTCGGAGCTGTCTGTCTT	809
Db	22201	CACGTCCTGAGCTGTGTCGCGGACCCGTCAACCCGTCAAGTGGTCTCGAAGGTCCCGCT	22142
Qy	810	CGGCAAGATCCGCGCGCGGGGCTTACGGCTTGGCTGACGGCGTCTCGGCGCCCTGGCGCGT	869
Db	22141	GGGACCGCTGCTTCCCGCGGCGCTGAGCTGTGGCGCGCTGACCGGCGCTGGCGCGCGAG	22082

RESULT 11
 US-08-474-933-1/C
 Sequence 1, Application US/08474933
 Patent No. 5866410
 GENERAL INFORMATION:
 APPLICANT: Ryan, Michael J.
 APPLICANT: Lotvin, Jason A.
 APPLICANT: Strachy, Nancy
 APPLICANT: Pantini, Susan E.
 TITLE OF INVENTION: Cloning of the biosynthetic pathway for
 TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmidas
 TITLE OF INVENTION: useful therein
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: American Cyanamid Company
 STREET: One Cyanamid Plaza
 CITY: Wayne
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07470
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/474,933
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/125,468
 FILING DATE: 22-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Tsevdos, Estelle J
 REGISTRATION NUMBER: 31,145
 REFERENCE/DOCKET NUMBER: 31,255-02
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (201) 831-3241
 TELEFAX: (201) 831-3305
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:

```

; LENGTH: 3001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-06-474-933-1

```

Query Match	4.8%	Score 51.2;	DB 2;	length 30001;
Best Local Similarity	44.2%;	Pred. No. 0.011;		
Matches 212;	Conservative 0;	Mismatches 268;	Indels 0;	Gaps 0;

Qy	390	CGAGCTCTCTTAAGCCCTTTTTCACCTGGAGCCCATGGGAGCATGCAAGCCTGATGGAGG	449
Db	22561	CAGAGAGCTCAAGGGATGTTCCGCTTCGCGCTCTGGAGCAGGACAGGAGGAATCCT	22502
Qy	450	CCGCAACCGTATCCCTGAAGCTGAGGCTGCGCTCTTCCGCAAGATGGCAACCGGCTT	509
Db	22501	GCTGATCCGGAGACGGATGGGAGCTCAAGCCGCTTACTACTGAGCCGACCGCCCGGCT	22442
Qy	510	GGGGCACTGTCACACAGACGAGTCTGGTCTCTGGGTGATCTCAAGCTGTATGCTTTGTCTT	569
Db	22441	GCGCTTCCGCTTCGGAACCCAGGCGCTCTTGGGGAGCCGAGACCTGTCCCGCGGGTCCG	22388
Qy	570	CGCTGACCGTGAGAGGAGAAGCTGTGCTGTGAGAACTTGAGGACTCTCTGCGTGTGAC	629
Db	22381	CCCGAGCGGCTGTGGGAGGTGCTTGACATGTGTGAAGACGCCGAGCGGCGGCTCTTCTC	22322
Qy	630	TGGGCGAGATGATTCCTCTGTGGGACAAAGCACGCGTCCAGGCTACTGTGGAGCTGAGAT	689
Db	22321	CGGACTCTGTGAGAGTTCGTCCGGGTCAAGTCTCTCCGGATGGGACCGGGGGATGCGCGG	22262
Qy	690	ACTCAGCTCACGGGGCTCATACTCGGGCAAGAGAGCGGATGTCTGGAGCTCTGGGCGTGGC	749
Db	22261	GCAGAGCTACTGGCGGGCTGGAAGGCCCGGCGACGACACGACGACCTGCGGACGACATCCG	22202
Qy	750	GCTCTTCAACCATGCTGACCGGCCACTTACCCCTTCCAGAGCTGAGAGCCTGTCTGTCTT	809
Db	22201	CACCGTCCGTAAGCTCTGCGCGGACACCGCTCACCCGTCAGCTGTGCTCGGAGTCCCGCT	22142
Qy	810	CGGCAAGATCCGCGCGGGGGGCTTAAGCCTTTGGCTGAGAGGCTCTTGCGCCCTTCCGCGCTG	869
Db	22141	GGGCAAGCTGCTTCTCGGCGGCGCTGAGTCTGTGCGCTGACGCGGCGCTGGCCGCGCGAG	22082

RESULT 12
US-09-691-861A-14
Sequence 14, Application US/09691861A
Patent No. 6482935
GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLED OF INVENTION:
FILE REFERENCE: CL000892
CURRENT APPLICATION NUMBER: US/09/691,861A
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 601
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(601)
OTHER INFORMATION: n = A,T,C or G
US-09-691-861A-14

Query Match 4.7%; Score 51; DB 4; Length 601;
Best Local Similarity 49.7%; Pred. No. 0.0062;
Matches 188; Conservative 0; Mismatches 181; Indels 9; Gaps 2;

420 CATTGGGAGCATGCAGCGCTGTGTGGAAAGCCGCCAACCGATTCCTGAGCCTGAAGCTGC 479

Db 73 CCAGGGGCACTCTCGAGTTAATCAAAACCGGGAGCCCTGCAATGAGACGAAGCTCG 132
Oy 480 CGTGTCTTCCGCGAGATGAGCACCCTGGCGCAGTGTCAACGACAGCGTGTGCT 539
Db 133 CAAAGAGTTCCACAGCTTCTTGGCCATCAAGTACAGCAGACCTGAGCTGTCCA 192
Oy 540 GCGTATCTCAAGCTGTGTGCTTGTCTTGTGCTGACCGTGAAGAAAGAACTGTGCT 599
Db 193 CCGGAGCTCAAGTGTGACAACTTCTCTTGACAAAGACTTCAACATCAAGCTGTCCGA 252
Oy 600 GGAAGAACTGAGAGCTCTGCGTGTGCTGAGCGGCAGATGATTCCTGTGTGGAAGAACA 659
Db 253 CTTCAGCTTCTCAAGCCCTGCTGCGGAGATGACAGTGTGCAATGCGNTTAAGCAAGC 312
Oy 660 C-----GCGTCCCGCAGCTAAGTGGAGCTGAGATGACTGAGCTCAGCGGCTCATATCTC 713
Db 313 CTTCTGTGGGTGACCAAGGATAGGGGCCCCAGAGGTGCT---GCAGGGGATTCCTTACCA 369
Oy 714 GGGCAGGACCGCATGTGTGTGAGCGTGGCGCTGCTTTCACCAATGTGGCCGCCA 773
Db 370 GCCCAAGGTGTACGACATCTGAGCGCTGATCTTCAATCATGATGTCTGCGGCTC 429
Oy 774 CTACCCCTTCAGGACTC 791
Db 430 CATGCCCTACGACGACTC 447

RESULT 13

US-09-691-861A-1
Sequence 1, Application US/09691861A
Patent No. 6482935

GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CLO00892
CURRENT APPLICATION NUMBER: US/09/691,861A
CURRENT FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1104
TYPE: DNA
ORGANISM: Homo sapiens
US-09-691-861A-1

Query Match 4.7%; Score 50.4; DB 4; Length 1104;
Best Local Similarity 49.7%; Pred. No. 0.0095;

Matches 188; Conservative 0; Mismatches 181; Indels 9; Gaps 2;

Oy 420 CCATGGGAGCATGCAACGCTGTGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGC 479
Db 282 CCAAGGGGCACTCTCGAGTTAATCAAAACCGGGAGCCCTGATGAGAGCAAGCTCG 341
Oy 480 CGTGTCTTCCGCGAGATGAGCACCCTGGCGCAGTGTCAACGACAGCTGTGCT 539
Db 342 CAAAGAGTTCCACAGCTTCTTGGCCATCAAGTACGCAAGCTGAGAGCTGTCCA 401
Oy 540 GCGTATCTCAAGCTGTGTGCTTGTCTTGTGCTGACCGTGAAGAAAGAACTGTGCT 599
Db 402 CCGGAGCTCAAGTGTGACAACTTCTCTTGACAAAGACTTCAACATCAAGCTGTCCGA 461
Oy 600 GGAAGAACTGAGAGCTCTGCGTGTGCTGAGCGGCAGATGATTCCTGTGTGGAAGAACA 659
Db 462 CTTCAGCTTCTCAAGCCCTGCTGCGGAGATGACAGTGTGCAATGAGCATTAAGCAAGC 521
Oy 660 C-----GCGTCCCGCAGCTAAGTGGAGCTGAGATGACTGAGCTCAGCGGCTCATATCTC 713
Db 522 CTTCTGTGGGTGACCAAGGATAGGGGCCCCAGAGGTGCT---GCAGGGGATTCCTTACCA 578
Oy 714 GGGCAGGACCGCATGTGTGTGAGCGTGGCGCTTTCACCAATGTGGCCGCCA 773

Db 579 GCCCAAGGTGTACGACATCTGAGCTTGAAGCGGTATCTTCAATCATGATGTCTGCGGCTC 638
Oy 774 CTACCCCTTCAGGACTC 791
Db 639 CATGCCCTACGACGACTC 656

RESULT 14

US-09-691-861A-3
Sequence 3, Application US/09691861A
Patent No. 6482935

GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CLO00892
CURRENT APPLICATION NUMBER: US/09/691,861A
CURRENT FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 9862
TYPE: DNA
ORGANISM: Homo sapiens
US-09-691-861A-3

Query Match 4.7%; Score 50.4; DB 4; Length 9862;
Best Local Similarity 49.7%; Pred. No. 0.014;

Matches 188; Conservative 0; Mismatches 181; Indels 9; Gaps 2;

Oy 420 CCATGGGAGCATGCAACGCTGTGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGC 479
Db 1784 CCAAGGGGCACTCTCGAGTTAATCAAAACCGGGAGCCCTGATGAGAGCAAGCTCG 1843
Oy 480 CGTGTCTTCCGCGAGATGAGCACCCTGGCGCAGTGTCAACGACAGCTGTGCT 539
Db 1844 CAAAGAGTTCCACAGCTTCTTGGCCATCAAGTACGCAAGCTGAGAGCTGTCCA 1903
Oy 540 GCGTATCTCAAGCTGTGTGCTTGTCTTGTGCTGACCGTGAAGAAAGAACTGTGCT 599
Db 1904 CCGGAGCTCAAGTGTGACAACTTCTCTTGACAAAGACTTCAACATCAAGCTGTCCGA 1963
Oy 600 GGAAGAACTGAGAGCTCTGCGTGTGCTGAGCGGCAGATGATTCCTGTGTGGAAGAACA 659
Db 1964 CTTCAGCTTCTCAAGCCCTGCTGCGGAGATGACAGTGTGCAATGAGCTTAAGCAAGC 2023
Oy 660 C-----GCGTCCCGCAGCTAAGTGGAGCTGAGATGACTGAGCTCAGCGGCTCATATCTC 713
Db 2024 CTTCTGTGGGTGACCAAGGATAGGGGCCCCAGAGGTGCT---GCAGGGGATTCCTTACCA 2080
Oy 714 GGGCAGGACCGCATGTGTGAGCGTGGCGCTGCTTTCACCAATGTGGCCGCCA 773
Db 2081 GCCCAAGGTGTACGACATCTGAGCGCTGAGCGTATCTTCAATCATGATGTCTGCGGCTC 2140
Oy 774 CTACCCCTTCAGGACTC 791
Db 2141 CATGCCCTACGACGACTC 2158

RESULT 15

US-09-016-434-536
Sequence 536, Application US/09016434
Patent No. 6500938

GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016.434
FILING DATE: HERSMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 536:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: NGANOT01
CLONE: 2314392
US-09-016-434-536

Query Match 4.6%; Score 49.4; DB 4; Length 263;
Best Local Similarity 47.3%; Pred. No. 0.012;
Matches 107; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY	727	GATGTCGAGACCTGGGCGCTCTTACACATGCTGGCCGACATACCCCTTCCAG	786
DB	10	GACATCTGAGCTGAGCTGATGATGACAGGCTGACTGACCCCACTTCATG	69
QY	787	GACTCGAGACCTGCTGCTGCTTCCGCAAGATCCGCGGAGCCCTACGCTTGCTGCA	846
DB	70	GCTTCACCCCTGNNNGAGATGTACCAAAACATCCGTTANGNCACATACCCGAGANCCT	129
QY	847	GAGCTCTGAGCCCTGCGCGCTGTCTGCTGCTGCTCTTCTGCTGGAGCCAGCTGAA	906
DB	130	CANCTGTGNCATGCTGCTGCTCATGTTCACTCTTANCAACCCANNNGNCAT	189
QY	907	CGGCTCACAGCACAAGGATCTCTGCAACCCCTGAGCTGCGACAGG	952
DB	190	CNNCCAGCCCTGAGCAGCTGAGGAGCACTTCTTCATACAGG	235

Search completed: January 16, 2004, 16:49:45
Job time : 72.5605 secs

Db 440 CTGGCCGCTGGAGCCCTTACGCGCGGCGCCCGCAGACAGATGCTGGCCACT 499
Qy 421 GAGGCTCTGGCTGATGACCAAGCTCTTACAGCCTTTTCACTCGAGCCATGAGGACATG 480
Db 500 GAGGCTCTGGCTGATGACCAAGCTCTTACAGCCTTTTCACTCGAGCCATGAGGACATG 559
Qy 481 CACAGCCTGATGACCAAGCCGCAAGATGCTTGAAGCCTGAGGCTGCTTTCCG 540
Db 560 CACAGCCTGATGACCAAGCCGCAAGATGCTTGAAGCCTGAGGCTGCTTTCCG 619
Qy 541 CAGATGACCAAGCCGCTGGCGCACTGTCACAGCAAGATGCTGATGCTGATGCTCAAG 600
Db 620 CAGATGACCAAGCCGCTGGCGCACTGTCACAGCAAGATGCTGATGCTGATGCTCAAG 679
Qy 601 CTGTGCTGCTTTGCTTCTGCTGACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 680 CTGTGCTGCTTTGCTTCTGCTGACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 739
Qy 661 GACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 740 GACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799
Qy 721 TACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 800 TACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 859
Qy 781 TGAAGCTGAGGCTGAGGCTGCTTACAGATGCTGAGGCTGAGGCTGAGGCTGAGGCTG 840
Db 860 TGAAGCTGAGGCTGAGGCTGCTTACAGATGCTGAGGCTGAGGCTGAGGCTGAGGCTG 919
Qy 841 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 920 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979
Qy 901 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 980 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1039
Qy 961 ACAGCAGAGAGATCTCTGCTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1006
Db 1040 ACAGCAGAGAGATCTCTGCTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1085

RESULT 2
US-09-220-132-135/c
Sequence 135, Application US/09220132
Patent No. 6506607

GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220.132
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25, 60/066,821
PRIOR APPLICATION NUMBER: US 60/066,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 135
LENGTH: 2559
TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-135

Query Match 24.5% Score 505.2; DB 4; Length 2559;
Best Local Similarity 94.6% Pred. No. 1.4e-99;
Matches 566; Conservative 0; Mismatches 28; Indels 4; Gaps 4;
Qy 604 TGTGCTTGTCTTCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663

Db 1634 TGACCTTCTGTTTCTCCCATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1575
Qy 664 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
Db 1574 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1516
Qy 724 GTGGAGCTGAGATCTAGCTCAAGGCTCTCTACTGAGGCTGAGGCTGAGGCTGAGG 783
Db 1515 GTGGAGCTGAGATCTAGCTCAAGGCTCTCTACTGAGGCTGAGGCTGAGGCTGAGG 1456
Qy 784 AGCTGAGGCTGAGGCTCTTCAAGATGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 843
Db 1455 AGCTGAGGCTGAGGCTCTTCAAGATGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1396
Qy 844 CCTGCTGCTGCTTCTGAGAGATGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 903
Db 1395 CCTGCTGCTGCTTCTGAGAGATGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1336
Qy 904 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
Db 1335 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1276
Qy 964 GCAAGAGATCTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023
Db 1275 G-CACAGAGATCTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
Qy 1024 GATTCATCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1083
Db 1217 GATTCATCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159
Qy 1084 AGGAG 1143
Db 1158 AGGAG 1099
Qy 1144 TCAGCTGCAAGAGATGATGTTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201
Db 1098 TCAGCTGCAAGAGATGATGTTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1041

RESULT 3
US-08-146-421-4/c
Sequence 4, Application US/08146421
Patent No. 5543499

GENERAL INFORMATION:
APPLICANT: BREMER, GARY
TITLE OF INVENTION: DNA SEQUENCE ENCODING A POLYPEPTIDE WITH
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSES: DILWORTH & BARRESE
STREET: 4350 LA JOLLA VILLAGE DRIVE, SUITE 300
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/146,421
FILING DATE: 29-OCT-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: PEPPER PH.D., FREDERICK W.
REGISTRATION NUMBER: 31,286
REFERENCE/DOCKET NUMBER: 489-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4410
TELEFAX: 619-453-2839
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:


```

APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSER: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1146:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1480860
US-09-016-434-1146

Query Match
Best Local Similarity 4.1%; Score 85; DB 4; Length 1302;
Matches 241; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 503 ACCGATCCCTGAGCTGAGCTGCTGCTCTTCCGCAAGATGCGACCGCTGAGCGC 562
DB 437 AGCGTTTCCAGTGTGCCAGGCCACGGGTACTTCTGTGAGCTGATGACGGCTGAGAT 496
QY 563 ACTGTACCAAGCAGGCTGTGCTGTGCTGATCTCAAGCTGTGCTGCTTGTCTTGCTG 622
DB 497 ACCTGATAGCCAGGCGCATGTGTGCAACAAGATCAAGCGGGGAACTGCTGTCAACA 556
QY 623 ACCGTGAGAGAAAGCTGTGTGTGAGAGAACTTGAGAGACTTCTGCTGTGACTGGGC 682
DB 557 CGGGTGGCACTCTCAAAATCTCCGACTGTGGCGTGGCCAGGACATGCAACCGCTTGGCGG 616
QY 683 CAGATGATTCCTGCTGGGCAAGACAGCGGTGCCAGGCCAAGTGGGAACTTGAGATCTCA 742
DB 617 CGGAGGACACCTGCGGACCAAGGAGCTCCCGGCTTTCAGCGCGCGGAGATTGCA 676
QY 743 GCTCAGGGGCTCATACTGTGCGCAAGGACCGGATGTGTGAGACCTGTGGCGCTGCT 802
DB 677 ACGGCTGGACACTTCTTCGGCTTCAAGGTGAGACATCTGTGTGGGTGACCCCTT 736
QY 803 TCACCATGTGGCGCGGCACTACCGCTTCAAGACTCGAGACTGTCTGTCTTGGCA 862
DB 737 ACAACATCAACAAGGCTGTGTACCCCTTGAAGGGGCAACAATCAAGATTGTTGAGA 796
QY 863 AGATCGCGCGGGGCTTGTGCTGTGCAAGGCTTCTGGGCGCGCTGCGCTGTGG 922

```

```

DB 797 ACATCGGAGAGGGAGCTAACGCCATCCGGGCGACTGTGGCCCCCGCTCTGTGACTGC 856
QY 923 TTCGCTGCTCTCTTCTGTGCGGAGCCAGCTGAAAGCGCTCAAGCCACAGGATCTCTGC 982
DB 857 TGAAGGATGCTTGTAGTACGAACCGGCGAAGAGGTTCTTCATCCGGGAGATCCGGCAGC 916
QY 983 ACCCTGGCTCGACAGACC 1003
DB 917 ACAGCTGGTCCGAAAGAAC 937

RESULT 6
US-09-559-023-1
Sequence 1, Application US/09559023
Patent No. 6551796
GENERAL INFORMATION:
APPLICANT: Abramson, Ruth
APPLICANT: Leal-Pinto, Edgar
APPLICANT: Lipkowitz, Michael
TITLE OF INVENTION: NUCLEIC ACID ENCODING URATE TRANSPORTER
FILE REFERENCE: 070165.0574
CURRENT APPLICATION NUMBER: US/09/559,023
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/221,898
PRIOR FILING DATE: 1998-12-28
PRIOR APPLICATION NUMBER: US 60/099,752
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: US 60/070,215
PRIOR FILING DATE: 1997-12-31
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 1
LENGTH: 1545
TYPE: DNA
ORGANISM: Rat
US-09-559-023-1

Query Match
Best Local Similarity 4.1%; Score 84.6; DB 4; Length 1545;
Matches 117; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1889 CTTGTGGGACACTTGGGCTCCACATCCAGCTCATCTAGGTTTGTGATACATG 1948
DB 1375 CTCGAGCTGACCCAGAAATGTCACATCTTTCCTCCATCTTCCCATCCCATAA 1434
QY 1949 AGTATGATGTTTACTCTGTCTATTAAGAGATTTATGAAATATATATATATATATAT 2008
DB 1435 ATAAAGAAATATCAAGCTTGTCTTACAAATATATATATATATATATATATATAT 1494
QY 2009 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2059
DB 1495 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1545

RESULT 7
US-09-996-243-300
Sequence 300, Application US/09996243
Patent No. 6478825
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Pong, Sherman
APPLICANT: Getber, Hanspeter
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Aublin L.

```

APPLICANT: KJjavin,Ivar J.
APPLICANT: Napier,Mary A.
APPLICANT: Pan,James
APPLICANT: Paoni,Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secrected and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P230P1C13
CURRENT APPLICATION NUMBER: US/09/996,243
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 4.0%; Score 82.8; DB 4; Length 1674;
Best Local Similarity 58.2%; Pred. No. 8.5e-09;
Matches 164; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

1779 TGAGGCTCCAGGCTGTCAACATGGGCTTCTGACCTGACACCAAGATTGAG-GGACA 1837
1368 TGAAGCTGCAAGATGCTGAGGAGGAGGCTACTTCCAGAGCTGTGGGCTCTGTCAACA 1427
1838 GGATTAGGACAGGCTGTCTGTCTGTGGCCACCTGGAAGTCCAGGTGGACTCTTCTGGGG 1897
1428 GCTTGCTGAAGAGCTGAGATGAGGCCCCACACCGAGACTGGAATGGATGCCCTAGTG 1487
1898 ACACCTGGGGTCCACATCCAGGTCTCACTAGTCTTTGGATATACATGATATGTAT 1957
1488 AGGCTGAGGGGTGACAGGCTGGGTGGCTTCTCAGGCGAGAGCATCTGGCAGTGTG 1547
1548 GCTTGCCATTAATGGAACCTGAAGCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1607
1548 GCTTGCCATTAATGGAACCTGAAGCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1607
2018 AA 2059
1608 AA 1649

RESULT 8
US-09-996-243-302
Sequence 302, Application US/09996243
Patent No. 6478825
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bostein, David
APPLICANT: Deenoyers, Luc

APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C13
CURRENT APPLICATION NUMBER: US/09/996,243
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05

1	PRIOR APPLICATION NUMBER: 60/0888212
2	PRIOR FILING DATE: 1998-06-05
3	PRIOR APPLICATION NUMBER: 60/0888217
4	PRIOR FILING DATE: 1998-06-05
5	PRIOR APPLICATION NUMBER: 60/0885555
6	PRIOR FILING DATE: 1998-06-09
7	PRIOR APPLICATION NUMBER: 60/0887344
8	PRIOR FILING DATE: 1998-06-10
9	PRIOR APPLICATION NUMBER: 60/0887389
10	PRIOR FILING DATE: 1998-06-10
11	PRIOR APPLICATION NUMBER: 60/0887422
12	PRIOR FILING DATE: 1998-06-10
13	PRIOR APPLICATION NUMBER: 60/0888100
14	PRIOR FILING DATE: 1998-06-10
15	PRIOR APPLICATION NUMBER: 60/0888242
16	PRIOR FILING DATE: 1998-06-10
17	PRIOR APPLICATION NUMBER: 60/0888262
18	PRIOR FILING DATE: 1998-06-10
19	PRIOR APPLICATION NUMBER: 60/0888588
20	PRIOR FILING DATE: 1998-06-11
21	PRIOR APPLICATION NUMBER: 60/0891053
22	PRIOR FILING DATE: 1998-06-12
23	PRIOR APPLICATION NUMBER: 60/0894400
24	PRIOR FILING DATE: 1998-06-16
25	PRIOR APPLICATION NUMBER: 60/0895122
26	PRIOR FILING DATE: 1998-06-17
27	PRIOR APPLICATION NUMBER: 60/0895144
28	PRIOR FILING DATE: 1998-06-16
29	PRIOR APPLICATION NUMBER: 60/0895232
30	PRIOR FILING DATE: 1998-06-17
31	PRIOR APPLICATION NUMBER: 60/0895388
32	PRIOR FILING DATE: 1998-06-17
33	PRIOR APPLICATION NUMBER: 60/0895999
34	PRIOR FILING DATE: 1998-06-17
35	PRIOR APPLICATION NUMBER: 60/0896000
36	PRIOR FILING DATE: 1998-06-17
37	PRIOR APPLICATION NUMBER: 60/0896533
38	PRIOR FILING DATE: 1998-06-17
39	PRIOR APPLICATION NUMBER: 60/0898011
40	PRIOR FILING DATE: 1998-06-18
41	PRIOR APPLICATION NUMBER: 60/0899077
42	PRIOR FILING DATE: 1998-06-18
43	PRIOR APPLICATION NUMBER: 60/0899088
44	PRIOR FILING DATE: 1998-06-18
45	PRIOR APPLICATION NUMBER: 60/0899477
46	PRIOR FILING DATE: 1998-06-19
47	PRIOR APPLICATION NUMBER: 60/0899498
48	PRIOR FILING DATE: 1998-06-19
49	PRIOR APPLICATION NUMBER: 60/0899552
50	PRIOR FILING DATE: 1998-06-19
51	PRIOR APPLICATION NUMBER: 60/0902466
52	PRIOR FILING DATE: 1998-06-22
53	PRIOR APPLICATION NUMBER: 60/0902522
54	PRIOR FILING DATE: 1998-06-22
55	PRIOR APPLICATION NUMBER: 60/0902544
56	PRIOR FILING DATE: 1998-06-22
57	PRIOR APPLICATION NUMBER: 60/0903499
58	PRIOR FILING DATE: 1998-06-23
59	PRIOR APPLICATION NUMBER: 60/0903555
60	PRIOR FILING DATE: 1998-06-23
61	PRIOR APPLICATION NUMBER: 60/0904299
62	PRIOR FILING DATE: 1998-06-24
63	PRIOR APPLICATION NUMBER: 60/0904311
64	PRIOR FILING DATE: 1998-06-24
65	PRIOR APPLICATION NUMBER: 60/0904355
66	PRIOR FILING DATE: 1998-06-24
67	PRIOR APPLICATION NUMBER: 60/0904444

```

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

```

Query Match	4.0%	Score 82.4;	DB 4;	Length 2136;
Best Local Similarity	85.2%;	Pred. No. 1.1e-08;		
Matches 92;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;

[illegible]

RESULT 9
 US-08-749-902-2
 ; Sequence 2, Application US/08749902
 ; Patent No. 5985635
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Goll, Susya K.
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: NOVEL HUMAN SRIN/THREONINE
 ; TITLE OF INVENTION: PROTEIN KINASES
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESSES:

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C13
CURRENT FILING DATE: 2001-11-14
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858

PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25

	Query Match	Best Local Similarity	89.5%; Matches 85;	Score 79.6; Conservative 0;	DB 4; Mismatches 10;	Length 1927; Indels 0;	Gaps 0;
OY	1965	TGNGCCATTAAGAGAAATTGAATATAAAAAAAAAAAAAAAAAAAAAA					2024
Dd	1766	TTTGCTTATAAGGAAATTCANNAAAAAAAAAAAAAAAAAAAAAAAA					1825
OY	2025	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA					2059
Dd	1826	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA					1860

RESULT 14
US-09-328

```

US-09-328-965-1
/ Sequence 1, Application US/09328965
/ Patent No. 6501008
/
/ GENERAL INFORMATION:
/ APPLICANT: Nevins, Donald J.
/ APPLICANT: Simmons, Carl
/ APPLICANT: The Regents of the University of California
/ TITLE OF INVENTION: Endo- and Exo-Glucanases and Gene
/ FILE REFERENCE: 023070-096600US
/ CURRENT APPLICATION NUMBER: US/09/328,965
/ CURRENT FILING DATE: 1999-06-09
/ EARLIER APPLICATION NUMBER: US 60/088,760
/ EARLIER FILING DATE: 1998-06-10
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 1091
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: maize coleoptile endo-1,3,1,4-beta glucanase cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (68)..(979)
/ OTHER INFORMATION: endo-1,3,1,4-beta glucanase
/ US-09-328-965-1

```

Query Match	3.8%	Score 79.2	DB 4	Length 1091
Best Local Similarity	75.0%	Pred. No. 4.4e+08		
Matches 99	Conservative 0	Mismatches 33	Indels 0	Gaps 0

Qy	1928	CTCTAGAGTTTGGATACCATGAGTATGTAATGTTTACCGTGGCTATATTAAGAGCAATAT	1987
Db	940	CTCTATGATTTCTTTTCTTTTCTTTTGTGTTTTATGATCGCATTAAGTTCAGTAGG	999
Qy	1988	GAATTAATAA	204
Db	1000	GCTAAA	1055
Qy	2048	AAAAAAAAAAAAA	2059
Db	1060	AAAAAAAAAAAAA	1071

RESULT 15
US-09-370-253-1
Sequence 1, Application US/09370253
Patent No. 6163792
GENERAL INFORMATION:
APPLICANT: Allen, Stephen M.
APPLICANT: Sakai, Hajime
APPLICANT: Thorpe, Catherine J.
TITLE OF INVENTION: Amino Acid Transporter
FILE REFERENCE: BB-1200
CURRENT APPLICATION NUMBER: US/09/370,253
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: 60/097,222
EARLIER FILING DATE: August 20, 1998
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Microsoft Office 97
SEQ ID NO 1
LENGTH: 1882
TYPE: DNA
ORGANISM: Hordeum vulgare
US-09-370-253-1

Query March	3.8%	Score 79.2	DB 3	Length 1882
Beet Local Similarity	83.3%	Pred. No. 5.2e+08		
Matches 90, Conservative	0	Mismatches 18	Indels 0	Gaps 0

1952 ATGTATGTTTACCTGTGCCTAATTAAGGAAATTATGAAATTAATAAAAAAAAAAAAAAAAAAAA 2011

Db	1760	AATGTAACGGGTCTTTTCTTCAGAAAAA	1819
Qy	2012	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2059
Db	1820	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1867

Search completed: January 16, 2004, 16:49:41
Job time : 135.44 BECS

Dh	395	CCTCTGGCTGCTCTCGGCGGTTCCCTGTTCAGAGAAAGACGGTTTGAGATTGATGACAC	454
Oy	121	TTAGATACGAGACGTCCTCCCTCAGAAATGAGCTCGAATGAGGCCCCAGCCCAACTGCC	180
Dh	455	TTAGATACGAGACGTCCTCCCTCAGAAATGAGCTCGAATGAGGCCCCAGCCCAACTGCC	514
Oy	181	CCCTGACCTGTTCCTCCCTGAGGCCCACTATCTCTCAGATTCCTTGACCTGTCTGTGGCACT	240
Dh	515	CCCTGACCTGTTCCTCCCTGAGGCCCACTATCTCTCAGATTCCTTGACCTGTCTGTGGCACT	574
Oy	241	GCCTCCCGCTTTCGAGCCCTATATGTCTCCCTGAGGCCGAGAGAGGCGGAGCCCTACCG	300
Dh	575	GCCTCCCGCTTTCGAGCCCTATATGTCTCCCTGAGGCCGAGAGAGGCGGAGCCCTACCG	634
Oy	301	GCCTTCACCTGACCCCTTACAGGCACTGAGATATACCTGCAAGGTGTACCCCTGCAGAGAAC	360
Dh	635	GCCTTCACCTGACCCCTTACAGGCACTGAGATATACCTGCAAGGTGTACCCCTGCAGAGAAC	694
Oy	361	CTGGCCGTGTCTGTGAGACCTTATGAGGCGGTGCGCCCGGCAAGAGATATGAGCTCTGGCCACT	420
Dh	695	CTGGCCGTGTCTGTGAGACCTTATGAGGCGGTGCGCCCGGCAAGAGATATGAGCTCTGGCCACT	754
Oy	421	GAGGCTCTGAGCTGTATACCCATCTCTACGCTCTTTTCACTCGGAGCCCATGAGGACATG	480
Dh	755	GAGGCTCTGAGCTGTATACCCATCTCTACGCTCTTTTCACTCGGAGCCCATGAGGACATG	814
Oy	481	CACAGCCTGTGTGAGAGCCGACCCGATATCCCTGAGCCTGAGGCTGCGGTCTTTCGCG	540
Dh	815	CACAGCCTGTGTGTGAGAGCCGACCCGATATCCCTGAGCCTGAGGCTGCGGTCTTTCGCG	874
Oy	541	CAGATGACCAACCGCCTTATGCGGCACTGTCAACAACAACGCTGTGTCTGTGATATCTCAG	600
Dh	875	CAGATGACCAACCGCCTTATGCGGCACTGTCAACAACAACGCTGTGTCTGTGATATCTCAG	934
Oy	601	CTGTGTGTGCTTTTGTCTTGTCTGTGACCGGTGAGAGAGAAAGCTGTGTGTGAGAACTGTGAG	660
Dh	935	CTGTGTGTGCTTTTGTCTTGTCTGTGACCGGTGAGAGAGAAAGCTGTGTGTGAGAACTGTGAG	994
Oy	661	GACTCTGTGTGTGTGACTGTGAGGACAGATGATTCCTGTGTGAGAACAGCAACGTGCCAGCC	720
Dh	995	GACTCTGTGTGTGTGACTGTGAGGACAGATGATTCCTGTGTGAGAACAGCAACGTGCCAGCC	1054
Oy	721	TACCTGTGAGACTGAGATATCTCAGCTCAACGAGCCTCATATCTGGGCGAAGGCAACCGATGTC	780
Dh	1055	TACCTGTGAGACTGAGATATCTCAGCTCAACGAGCCTCATATCTGGGCGAAGGCAACCGATGTC	1114
Oy	781	TGAGAGCCTGTGTGTGTGTCTTCTTCAACATGTGAGGCGGCGCACTAACCCCTTCAAGATCTG	840
Dh	1115	TGAGAGCCTGTGTGTGTGTCTTCTTCAACATGTGAGGCGGCGCACTAACCCCTTCAAGATCTG	1174
Oy	841	GAGCCTGTCTGTCTTCTGTGAGAGATTCGAGCGGCGGAGCTTACGCTTGTCTGTGAGGCTCTC	900
Dh	1175	GAGCCTGTCTGTCTTCTGTGAGAGATTCGAGCGGCGGAGCTTACGCTTGTCTGTGAGGCTCTC	1234
Oy	901	TGGGCCCCCTGCGCTGTCTGTGTGTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	960
Dh	1235	TGGGCCCCCTGCGCTGTCTGTGTGTGTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1294
Oy	961	ACAGCGAAGAGCAATCTCTCTGCAACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1020
Dh	1295	ACAGCGAAGAGCAATCTCTCTGCAACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1354
Oy	1021	ACCGGATCCCATCTCTGTGAGAGGCTGTCCAGAGGTGTCTCTGTATGTGACTGTGAGCGAA	1080
Dh	1355	ACCGGATCCCATCTCTGTGAGAGGCTGTCCAGAGGTGTCTCTGTATGTGACTGTGAGCGAA	1414
Oy	1081	GCAGAGGAAAGAGAGAGAGACAGAGAGGTGTCTGTATGTGTGAGACCAACCTTACTACA	1140
Dh	1415	GCAGAGGAAAGAGAGAGAGACAGAGAGGTGTCTGTATGTGTGAGACCAACCTTACTACA	1474
Oy	1141	CGCTCACTGTCTCAACATGTGAGATGTGAGTTTGTGGGGGTAGCTTCAAGGCTTCTCTGTGCTG	1200
Dh	1475	CGCTCACTGTCTCAACATGTGAGATGTGAGTTTGTGGGGGTAGCTTCAAGGCTTCTCTGTGCTG	1534

OY	1201	AACGAGCCAAACCCTTAGGGCCTTCCAGAAAGGAGAAAGCAAGAACCTGTGTGGAGTG	1260
Db	1535	AACGAGCCAAACCCTTAGGGCCTTCCAGAAAGGAGAAAGCAAGAACCTGTGTGGAGTG	1594
OY	1261	TGCTGTGTACAATCTGTGCTTTGTTTCCAACAACAATGCAGTTCTGTCTTGGGTGCTTAAC	1320
Db	1595	TGCTGTGTACAATCTGTGCTTTGTTTCCAACAACAATGCAGTTCTGTCTTGGGTGCTTAAC	1654
OY	1321	GTCGAAGCCCTGTCTTCGGGTGCTGTGGAGATACAGCATGTGAGCAAAGAGCAATATTCCC	1380
Db	1655	GTCGAAGCCCTGTCTTCGGGTGCTGTGGAGATACAGCATGTGAGCAAAGAGCAATATTCCC	1714
OY	1381	TGCTCACAGAGATGACAAACTGGGATGCTTGAAGCTGACAAACATTTTCCATGACATAGG	1440
Db	1715	TGCTCACAGAGATGACAAACTGGGATGCTTGAAGCTGACAAACATTTTCCATGACATAGG	1774
OY	1441	TCACTGTCTACACTGGGTACCTTTTGTACCAAGTGTGGGCTCCACTGATGCTGTGCTCA	1500
Db	1775	TCACTGTCTACACTGGGTACCTTTTGTACCAAGTGTGGGCTCCACTGATGCTGTGCTCA	1834
OY	1501	GGCACCTCTGTCCAAAGGACATCCCTTTCACAAACAAACAGCGTCCCTTGTATCTTGTA	1560
Db	1835	GGCACCTCTGTCCAAAGGACATCCCTTTCACAAACAAACAGCGTCCCTTGTATCTTGTA	1894
OY	1561	CCTTTTCAGAGAAAAGGAGGTATCCCTGTGCCAAAGGCTCCAGGCTCTCCCTGCAACT	1620
Db	1895	CCTTTTCAGAGAAAAGGAGGTATCCCTGTGCCAAAGGCTCCAGGCTCTCCCTGCAACT	1954
OY	1621	CAGGACCCAAAGCCCACTCACTCTGGGAACTGTGTCCCAAGACATCTGTCTCTTGATT	1680
Db	1955	CAGGACCCAAAGCCCACTCACTCTGGGAACTGTGTCCCAAGACATCTGTCTCTTGATT	2014
OY	1681	AAGAGATTCTCTTTCAGAGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAACTA	1740
Db	2015	AAGAGATTCTCTTTCAGAGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAACTA	2074
OY	1741	TGAGGCTAAGTTCTGTCTTAACCTCAAGACTGTTCGTGAAATGAGGGTCCAGGCCCTGTCAAC	1800
Db	2075	TGAGGCTAAGTTCTGTCTTAACCTCAAGACTGTTCGTGAAATGAGGGTCCAGGCCCTGTCAAC	2134
OY	1801	ATGGGGCTTCTTGACCTGAGACCAAGGTTGAGGGAACGAGATTAGGCAAGGCTGTCTGT	1860
Db	2135	ATGGGGCTTCTTGACCTGAGACCAAGGTTGAGGGAACGAGATTAGGCAAGGCTGTCTGT	2194
OY	1861	GGCCACCTGGGAAGCTCCCAAGTGGGACTTTCTGGGGACACTTGGGGTCCACAATCCCAG	1920
Db	2195	GGCCACCTGGGAAGCTCCCAAGTGGGACTTTCTGGGGACACTTGGGGTCCACAATCCCAG	2254
OY	1921	GTCGATACCTTAGGTTTGGATTACAGATATGATATGTTTACCTGTGGCTAATTAAGA	1980
Db	2255	GTCGATACCTTAGGTTTGGATTACAGATATGATATGTTTACCTGTGGCTAATTAAGA	2314
OY	1981	GAATTATCAAT	2040
Db	2315	GAATTATCAAT	2374
OY	2041	AAAAAAAAAAAAAA 2055	
Db	2375	AAAAAAAAAAAAAA 2389	

RESULT 2
US-10-098-841-271

US-10-098-841-271

Publication No. US20020197679A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: LIU, Chienling

APPLICANT: Xu, Chongjun

APPLICANT: Zhou, Ping

! APPLICANT: Ma, Jungling

DB 1711 CAGAACCCAGCCAGCTCACTGAGGAACTGTGTTCCAGACATCTGTCTTGAAT 1770
 QY 1681 AAGAGATTCTCTTCAGAGCCCTTAAGCCCTGAGATTGGGCCAGAGATTAAGATCTCAAATA 1740
 DB 1771 AAGAGATTCTCTTCAGAGCCCTTAAGCCCTGAGATTGGGCCAGAGATTAAGATCTCAAATA 1830
 QY 1741 TGAAGCTAGTTCTTGTCTTAATCTCAAGACTGTGTTCTGAGGATGAGGGCTCAAGCCCTGCAACC 1800
 DB 1831 TGAGGCTAGTTCTTGTCTTAATCTCAAGACTGTGTTCTGAGGATGAGGGCTCAAGCCCTGCAACC 1890
 QY 1801 ATGGGAGCTTCTGAGCCCTGAGCCAGAGGTTGAGGAGACAGAGATTAGGAGGCTGTCTCTGT 1860
 DB 1891 ATGGGAGCTTCTGAGCCCTGAGCCAGAGGTTGAGGAGACAGAGATTAGGAGGCTGTCTCTGT 1950
 QY 1861 GGCACCTGAGGAAAGTCCAGAGTGGAGACTTTCTGAGGACACTTGGAGCTCAAGATCCAG 1920
 DB 1951 GGCACCTGAGGAAAGTCCAGAGTGGAGACTTTCTGAGGACACTTGGAGCTCAAGATCCAG 2010
 QY 1921 GTCCATCTCTAGGTTTGTGATTAACCATGAGTATGTTTACCTGAGCTAATAAGGA 1980
 DB 2011 GTCCATCTCTAGGTTTGTGATTAACCATGAGTATGTTTACCTGAGCTAATAAGGA 2070
 QY 1981 GAAATTATGAAATTAATAAAAAA 2002
 DB 2071 GAAATTATGAAATTAATAAAAAA 2092

RESULT 3
 US-09-925-301-260
 / Sequence 260, Application US/09925301
 / Patent No. US2002005308A1
 / GENERAL INFORMATION:
 / APPLICANT: Rosen et al.
 / TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 / PILE REFERENCES: P106
 / CURRENT APPLICATION NUMBER: US/09/925,301
 / PRIOR FILING DATE: 2001-08-10
 / PRIOR APPLICATION NUMBER: PCT/US00/05882
 / PRIOR FILING DATE: 2000-03-08
 / PRIOR APPLICATION NUMBER: 60/124,270
 / PRIOR FILING DATE: 1999-03-12
 / NUMBER OF SEQ ID NOS: 1694
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 260
 / LENGTH: 2048
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURES:
 / NAME/KEY: misc_feature
 / LOCATION: (66)
 / OTHER INFORMATION: n equals a, t, g, or c
 / NAME/KEY: misc_feature
 / LOCATION: (67)
 / OTHER INFORMATION: n equals a, t, b, or c
 / US-09-925-301-260

Query Match 85.9%; Score 1767.8; DB 9; Length 2048;
 Best Local Similarity 95.9%; Pred. No. 0;
 Matches 1882; Conservative 4; Mismatches 1; Indels 76; Gaps 4;

DB 305 -----GAGCCCGAGAGGCGAG 322
 QY 289 CCGGCTTACCGGAGCCCTGCACTGCGCTTACAGGCACTGAGTATTAAGTGAAGGTATACCC 348
 DB 323 CCGGCTTACAGAG-CCTGCACTGCGCTTACAGGCACTGAGTATTAAGTGAAGGTATACCC 380
 QY 349 GTCCAGGAAAGCCCTGAGCCCTGCTGAGAGCCCTTACCGCGAGCTGCCCCGACAGCATGTG 408
 DB 381 GTCCAGGAAAGCCCTGAGCCCTGCTGAGAGCCCTTACCGCGAGCTGCCCCGACAGCATGTG 440
 QY 409 GCTGAGCCCACTGAGAGCTCTGATGTAACCAAGCTCTCTACAGCTTTTTCACCTGGACC 468
 DB 441 GCTGAGCCCACTGAGAGCTCTGATGTAACCAAGCTCTCTACAGCTTTTTCACCTGGACC 500
 QY 469 CATGGAGGACATGACAGAGCTGAGGAGAGCCGACAGGATCCCTGAGAGCTGAGCTG 528
 DB 501 CATGGAGGACATGACAGAGCTGAGGAGAG-CGACACCTTATCTGAGCTGAGAGCTG 559
 QY 529 GTGCTCTTCCGCAAGATGAGCCAGCCGCTGAGGCACTGTACACAGACAGGCTGTGCTG 588
 DB 560 GTGCTCTTCCGCAAGATGAGCCAGCCGCTGAGGCACTGTACACAGACAGGCTGTGCTG 619
 QY 589 GGTGATCTCAAGCTGTGTGCTTTGTCTTCTGCTGACCCGTAAGAGAGAGAGGTGCTG 648
 DB 620 GGTGATCTCAAGCTGTGTGCTTTGTCTTCTGCTGACCCGTAAGAGAGAGAGGTGCTG 679
 QY 649 GAGAACTGAGAGGACCTCCGAGCTGAGCTGAGGAGGAGATTCCTGTGAGGACAGAC 708
 DB 680 GAGAACTGAGAGGACCTCCGAGCTGAGCTGAGGAGGAGATTCCTGTGAGGAGAGAC 739
 QY 709 GCGTCCAGGCTTACCTGAGGAGCTGAGATTAAGTCAAGGAGCTCATATCTGGGCAAG 768
 DB 740 GCGTCCAGGCTTACCTGAGGAGCTGAGATTAAGTCAAGGAGGCTCATATCTGGGCAAG 799
 QY 769 GAGAGGATGTGAGAGCTGAGGAGGAGGAGCTTACAGATGCTGAGGAGGAGGAGGAG 828
 DB 800 GAGAGGATGTGAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 859
 QY 829 TTCAGAGACTGAGAGGCTGCTGCTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 888
 DB 860 TTCAGAGACTGAGAGGCTGCTGCTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 919
 QY 889 CCTGAGGAGCTTGGAGCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 948
 DB 920 CCTGAGGAGCTTGGAGCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 979
 QY 949 GCTGAAAGGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1008
 DB 980 GCTGAAAGGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1039
 QY 1009 CCTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1068
 DB 1040 CCTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
 QY 1069 GAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1128
 DB 1100 GAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1159
 QY 1129 GAGCTTACCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1188
 DB 1160 GAGCTTACCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1219
 QY 1189 GTCTGAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1248
 DB 1220 GTCTGAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1279
 QY 1249 GTGTGAGAGT 1308
 DB 1280 GTGTGAGAGT 1339
 QY 1309 GTGTGAGAGT 1368

Db 395 CCTGCTGCTGCTCTGCGGGTTCCCTGTCAGAGAAAGCGTTGAGATTGATACAC 454
 Qy 121 TTAGATACGAGCGTCCGCTCCAGAAAGAGCTCGAAGTGGGCCCCAGGCCAGACTGCC 180
 Db 455 TTAGATACGAGCGTCCGCTCCAGAAAGAGCTCGAAGTGGGCCCCAGGCCAGACTGCC 514
 Qy 181 CCGTGTCTGTGGCCCTGAGCCCACTACTGCTCCAGATTCGAGCACTGTGTGGCCACT 240
 Db 515 CCGTGTCTGTGGCCCTGAGCCCACTACTGCTCCAGATTCGAGCACTGTGTGGCCACT 574
 Qy 241 GCGTCCGCTGTGGCCCTGATGTCTCTGAGAGCCGAGAGAGGGGGGCGGACTACCGG 300
 Db 575 GCGTCCGCTGTGGCCCTGATGTCTCTGAGAGCCGAGAGAGGGGGGCGGACTACCGG 634
 Qy 301 GCGTGTGACCTGCTTGGAGCCCTTATGTCTCTGAGAGCCGAGAGAGGGGGGCGGACT 360
 Db 635 GCGTGTGACCTGCTTGGAGCCCTTATGTCTCTGAGAGCCGAGAGAGGGGGGCGGACT 694
 Qy 361 CTTGCGCGTGTGGAGCCCTTATGGGGGCTGCGCCGCAAGAGATGTGGCTGGGCCACT 420
 Db 695 CTTGCGCGTGTGGAGCCCTTATGGGGGCTGCGCCGCAAGAGATGTGGCTGGGCCACT 754
 Qy 421 GAGGTCTGTGGCTGTGACCAAGCTCTCTACGCGCTTTTCACTCGAGCCATGGGAGACATG 480
 Db 755 GAGGTCTGTGGCTGTGACCAAGCTCTCTACGCGCTTTTCACTCGAGCCATGGGAGACATG 814
 Qy 481 CACAGCTGTGTGGAGAGCCGCAAGCTATCCCTGAGCCCTGAGCTGTGGCTGTGGCC 540
 Db 815 CACAGCTGTGTGGAGAGCCGCAAGCTATCCCTGAGCCCTGAGCTGTGGCTGTGGCC 874
 Qy 541 CAGATGTGGAGAGCCGCTGAGCGCACTGTACAGAGAGGCTGTGGCTGTGGCTGTGGAG 600
 Db 875 CAGATGTGGAGAGCCGCTGAGCGCACTGTACAGAGAGGCTGTGGCTGTGGCTGTGGAG 934
 Qy 601 CTGTGTGCTGTGTGTCTGTGTGACCGGTGAGAGAGAAAGCTGTGTGTGTGTGTGTGTGT 660
 Db 935 CTGTGTGCTGTGTGTCTGTGTGACCGGTGAGAGAGAAAGCTGTGTGTGTGTGTGTGTGT 994
 Qy 661 GACTCTGT 720
 Db 995 GACTCTGT 1054
 Qy 721 TACGTGTGGAGCTGAGATACTGAGCTACAGGAGCTTACTGAGGAGAGGAGCGAGTGT 780
 Db 1055 TACGTGTGGAGCTGAGATACTGAGCTACAGGAGCTTACTGAGGAGAGGAGCGAGTGT 1114
 Qy 781 TGAAGCTGTGGAGCTGT 840
 Db 1115 TGAAGCTGTGGAGCTGT 1174
 Qy 841 GAGCTGT 900
 Db 1175 GAGCTGT 1234
 Qy 901 TCGGCGCTGTGGCGCTGT 960
 Db 1235 TCGGCGCTGTGGCGCTGT 1294
 Qy 961 ACAGCCAGAGGAGCTCTCTGTGACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
 Db 1295 ACAGCCAGAGGAGCTCTCTGTGACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1354
 Qy 1021 ACCCGATGCCATTTCTGTGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 Db 1355 ACCCGATGCCATTTCTGTGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1414
 Qy 1081 GCCAG 1140
 Db 1415 GCCAG 1474
 Qy 1141 CGCTCAGCTGTGCAACAGT 1200
 Db 1475 CGCTCAGCTGTGCAACAGT 1534

Qy 1201 AACTGAGCCAAACCTTCAGTGTCTTCCAGAAAGGAGAAAGGAGAGAGCTGTGTGTGTGT 1260
 Db 1535 AACTGAGCCAAACCTTCAGTGTCTTCCAGAAAGGAGAAAGGAGAGAGCTGTGTGTGTGT 1594
 Qy 1261 TGT 1320
 Db 1595 TGT 1654
 Qy 1321 GTGCAAGCCCTGT 1380
 Db 1655 GTGCAAGCCCTGT 1714
 Qy 1381 TGT 1440
 Db 1715 TGT 1774
 Qy 1441 TCACTGT 1500
 Db 1775 TCACTGT 1834
 Qy 1501 GGCACCTGT 1560
 Db 1835 GGCACCTGT 1894
 Qy 1561 CTTTTTCAAG 1620
 Db 1895 CTTTTTCAAG 1954
 Qy 1621 CAGAGCCCAAGCCCAAGCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
 Db 1955 CAGAGCCCAAGCCCAAGCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2014
 Qy 1681 AAGAGATTCCTCTTCCAGGCTTAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
 Db 2015 AAGAGATTCCTCTTCCAGGCTTAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2074
 Qy 1741 TGAAGCTGT 1800
 Db 2075 TGAAGCTGT 2134
 Qy 1801 ATGGGGCTTCTGACCTGTGAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
 Db 2135 ATGGGGCTTCTGACCTGTGAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2194
 Qy 1861 GGCACCTGTGAAAGTCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
 Db 2195 GGCACCTGTGAAAGTCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2254
 Qy 1921 GTCCATCTGTAGGTTTGT 1980
 Db 2255 GTCCATCTGTAGGTTTGT 2314
 Qy 1981 GAAATTTATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2040
 Db 2315 GAAATTTATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2374
 Qy 2041 AAAAAAAAAAAAAAAAAA 2055
 Db 2375 AAAAAAAAAAAAAAAAAA 2389

RESULT 2
 US-10-098-841-271
 ; Sequence 271, Application US/10098841
 ; Publication No. US20020197679A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Xu, Chongjun
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yungqing

Db 1711 CAGAGCCCAAGCCAGCTCCTGGAACCTGTGTCCAGACATCTGTCTTGAATT 1770
Qy 1681 AAGAGATTCTCTCTCCAGGCTTAAGCTTGGAATTGGCCAGAGATAAATCCAACTA 1740
Db 1771 AAGAGATTCTCTCTCCAGGCTTAAGCTTGGAATTGGCCAGAGATAAATCCAACTA 1830
Qy 1741 TGAGGCTAGTTCTGTCTTACTCAAGACTGTCTTGGAATTGAGGGTCCAGGCTGTCAAC 1800
Db 1831 TGAGGCTAGTTCTGTCTTACTCAAGACTGTCTTGGAATTGAGGGTCCAGGCTGTCAAC 1890
Qy 1801 ATGGGGCTTCTGACCTGAGCAACAGTTGAGGAGCAGAGATTAGGCAAGGCTGTCTGT 1860
Db 1891 ATGGGGCTTCTGACCTGAGCAACAGTTGAGGAGCAGAGATTAGGCAAGGCTGTCTGT 1950
Qy 1861 GGCCACCTGAAAGTCCAGGCTGGAGCTCTTCTGGGAGCACTTGGGGTCCAAATCCAG 1920
Db 1951 GGCCACCTGAAAGTCCAGGCTGGAGCTCTTCTGGGAGCACTTGGGGTCCAAATCCAG 2010
Qy 1921 GTCCATCTCTAGGTTTGGATACATAGATATGTAATGTTAACTGTCCTAATAAAGA 1980
Db 2011 GTCCATCTCTAGGTTTGGATACATAGATATGTAATGTTAACTGTCCTAATAAAGA 2070
Qy 1981 GAATTATGAATAAAAAAAAA 2002
Db 2071 GAATTATGAATAAAAAAAAA 2092

RESULT 3
US-09-925-301-260
Sequence 260, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 260
LENGTH: 2048
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc_feature
LOCATION: (66)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (67)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-260

Query Match 85.9%; Score 1767.8; DB 9; Length 2048;
Bee Local Similarity 95.9%; Freq. No. 0;
Matches 1882; Conservative 4; Mismatches 1; Indels 76; Gaps 4;

Qy 49 ATGGAGGCAACCCCTGAGCTGCTCTGCGGGTTCCCTGACAGAAAGACGGTTGAG 108
Db 156 ATGGAGGCAACCCCTGAGCTGCTCTGCGGGTTCCCTGACAGAAAGACGGTTGAG 215
Qy 109 TTGATGACAACTTATAGATACGAGGCTCCGTCAGAAACGAGCTCGAAGTGGCCCCAG 168
Db 216 TTGATGACAACTTATAGATACGAGGCTCCGTCAGAAACGAGCTCGAAGTGGCCCCAG 275
Qy 169 CCCAGAGTGGCCCCCTGCTGTGCGCCCTGAGCCCACTACTGCTCCAGATCGTCAACT 228
Db 276 CCCAGAGTGGCCCCCTGCTGTGCGCCCTGAGCCCACTACTGCTCCAGATCGTCAACT 304
Qy 229 GCTGTGACCACTGCTCCGCTCTTGGGCCCTATGTCTCTCTGAGCCCGAGAGAGGCGGG 288

Db 305 -----GAGCCCGAGAGGCGGG 322
Qy 289 CGGGCTAACCGGGCCCTGCACTGCTCTACAGGCACTGATTAATCTGCAAGGTGTACCC 348
Db 323 CGGGCTAACCGGGCCCTGCACTGCTCTACAGGCACTGATTAATCTGCAAGGTGTACCC 380
Qy 349 GTCCAGAAAGCCCTGAGCCGTGTGAGGCTTACCGCGGAGCTGCCCCGACAGAGATGTG 408
Db 381 GTCCAGAAAGCCCTGAGCCGTGTGAGGCTTACCGCGGAGCTGCCCCGACAGAGATGTG 440
Qy 409 GCTGGGCCCACTGAGGCTCTGAGCTGTATACCAAGTCTCTTACCTTCTGACCTGAGCC 468
Db 441 GCTGGGCCCACTGAGGCTCTGAGCTGTATACCAAGTCTCTTACCTTCTGAGCC 500
Qy 469 CATGGAGACATGACACACCTGAGTGGAGGCGGACACGATATCCCTGAGCCCTGAGGCTG 528
Db 501 CATGGAGACATGACACACCTGAGTGGAGGCGGACACGATATCCCTGAGCCCTGAGGCTG 559
Qy 529 GTGCTCTTCCGACAGATGAGCCACCGCCCTGAGGCACTGTACACAGACAGTCTGCTGTG 588
Db 560 GTGCTCTTCCGACAGATGAGCCACCGCCCTGAGGCACTGTACACAGACAGTCTGCTGTG 619
Qy 589 CTGATCTCAAGCTGTGCTGCTTGTGTGCTGACCGGTGAGAGAAAGCTGCTGTG 648
Db 620 CTGATCTCAAGCTGTGCTGCTTGTGTGCTGACCGGTGAGAGAAAGCTGCTGTG 679
Qy 649 GAGAACCTGAGAGACTCTGCTGCTGCTGACCTGAGGCAATGATTCCTGTGAGAACAGAC 708
Db 680 GAGAACCTGAGAGACTCTGCTGCTGCTGACCTGAGGCAATGATTCCTGTGAGAACAGAC 739
Qy 709 GCGTCCCAAGCTTACGTGGAGCTGAGATACAGCTCACGGGCTCTACTGCGGCAAG 768
Db 740 GCGTCCCAAGCTTACGTGGAGCTGAGATACAGCTCACGGGCTCTACTGCGGCAAG 799
Qy 769 GCAGCCGATGTGAGAGCTGAGGCTGAGGCTCTTACCAAGCTGAGGCGGCACTAACCC 828
Db 800 GCAGCCGATGTGAGAGCTGAGGCTGAGGCTCTTACCAAGCTGAGGCGGCACTAACCC 859
Qy 829 TTCCAGAGCTCGAGGCTGTCTCTGCTTCTGCGCAAGATCCGCGCGGAGCTTACGCTT 888
Db 860 TTCCAGAGCTCGAGGCTGTCTCTGCTTCTGCGCAAGATCCGCGCGGAGCTTACGCTT 919
Qy 889 CTGCAAGGCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
Db 920 CTGCAAGGCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979
Qy 949 GCTGAAGGCTCACAGGCAACAGGATCTCTCTGACCCCTGAGCTGAGCAAGACCCGATG 1008
Db 980 GCTGAAGGCTCACAGGCAACAGGATCTCTCTGACCCCTGAGCTGAGCAAGACCCGATG 1039
Qy 1009 CCTTAAAGCCCAACCCGATCCATCTCTGAGAGGCTGCCAGAGTGTCTCTGATGAGACTG 1068
Db 1040 CCTTAAAGCCCAACCCGATCCATCTCTGAGAGGCTGCCAGAGTGTCTCTGATGAGACTG 1099
Qy 1069 GGGCTGACAGAGCTGAGGAAAGAGAGAGAGACAGAGAGTGTCTGATGAGTGTGAGAC 1128
Db 1100 GGGCTGACAGAGCTGAGGAAAGAGAGAGAGACAGAGAGTGTCTGATGAGTGTGAGAC 1159
Qy 1129 CACCCTTACAGGCTCAGGCTGAGGCAACAGTGAATTGAGGAGTGTGAGGCTT 1188
Db 1160 CACCCTTACAGGCTCAGGCTGAGGCAACAGTGAATTGAGGAGTGTGAGGCTT 1219
Qy 1189 CTCCTGCTCTGAACTGAGGCAACCTTCAAGTCTCTTCAAGAGGAGAAAGGAGAGAC 1248
Db 1220 CTCCTGCTCTGAACTGAGGCAACCTTCAAGTCTCTTCAAGAGGAGAAAGGAGAGAC 1279
Qy 1249 CTGTGTGAGATGTGTGTGTATACATCTCTTGTTCACACACATGAGTCTGTCTTG 1308
Db 1280 CTGTGTGAGATGTGTGTGTATACATCTCTTGTTCACACACATGAGTCTGTCTTG 1339
Qy 1309 GGTGCTTATCAGAGTGCAGAGCCCTGTTCTGAGTGTGAGGTACAGAGTGTGAGCAAGGA 1368

Db 1340 GGTCTTATCAGTGTCCAAAGCCCTGTTCTCGGTGTGAGTACAGCAGTGAAGAAAGA 1399
Qy 1369 GACAAATATTCCTCTGTCAAGAGATGACAACTGGCAATCTTGTAGCTGACAACTTTTC 1428
Db 1400 GACAAATATTCCTCTGTCAAGAGATGACAACTGGCAATCTTGTAGCTGACAACTTTTC 1459
Qy 1429 CATGACCAATAGTCACTGTCTACACTGGGTACACTTGTGTACAGTGTGGCTCACTGA 1488
Db 1460 CATACCAATAGTCACTGTCTACACTGGGTACACTTGTGTACAGTGTGGCTCACTGA 1519
Qy 1489 TGTGTGTGTCTGAGCACTCTGTGTCCAAAGACAAATCTTGTCAAAACAAACAGTGTCT 1548
Db 1520 TGTGTGTGTCTGAGCACTCTGTGTCCAAAGACAAATCTTGTCAAAACAAACAGTGTCT 1579
Qy 1549 TTTGATCTTGTACTCTTTTCAAGAAAGAGATATCTCTGTGCCAAAGGCTTCCAGCTCTC 1608
Db 1580 TTTGATCTTGTACTCTTTTCAAGAAAGAGATATCTCTGTGCCAAAGGCTTCCAGCTCTC 1639
Qy 1609 TCCCTGTCAACTCAGAACCCCAAGGCTCACTGTGGAACTGTGTCCAGCACTCTCT 1668
Db 1640 TCCCTGTCAACTCAGAACCCCAAGGCTCACTGTGGAACTGTGTCCAGCACTCTCT 1699
Qy 1669 GTCTCTTGTATTAAGAGATTCTCTTCCAGGCTTAAGCTTGGAAATTTGGCCCAAGATTA 1728
Db 1700 GTCTCTTGTATTAAGAGATTCTCTTCCAGGCTTAAGCTTGGAAATTTGGCCCAAGATTA 1759
Qy 1729 GAATCCAACTATAGAGCTAGTCTTGTCTTAACCTCAAGACTTGTCTGAAATGAGAGTCCA 1788
Db 1760 GAATCCAACTATAGAGCTAGTCTTGTCTTAACCTCAAGACTTGTCTGAAATGAGAGTCCA 1819
Qy 1789 GGCCTGTCAACCATAGGGCTTCTGACCTGACCAACCAAGTTGAGGACAGAGATTAGCAG 1848
Db 1820 GGCCTGTCAACCATAGGGCTTCTGACCTGACCAACCAAGTTGAGGACAGAGATTAGCAG 1879
Qy 1849 GGTGTGTCTGTGGCCCACTGGAAGTCCCAAGTGGGACTTTCTGTGGGACACTTGGGCT 1908
Db 1880 GGTGTGTCTGTGGCCCACTGGAAGTCCCAAGTGGGACTTTCTGTGGGACACTTGGGCT 1939
Qy 1909 CCACAATCCCAAGGTCATCTCTAGGTTTGTGATACCAATGATATGATTTTACTGTG 1968
Db 1940 CCACAATCCCAAGGTCATCTCTAGGTTTGTGATACCAATGATATGATTTTACTGTG 1999
Qy 1969 CCAATTAAGAGAGATTATGAAATTAATAAAAAAAAAAAAAA 2011
Db 2000 CCAATTAAGAGAGATTATGAAATTAATAAAAAAAAAAAAAA 2042

RESULT 4
US-09-799-875-9
/ Sequence 9, Application US/09799875
/ Patent No. US20020034780A1
/ GENERAL INFORMATION:
/ APPLICANT: Meyers, Rachel
/ APPLICANT: Kapeller-Libermann, Rosana
/ TITLE OF INVENTION: No. US20020034780A1 Human Protein Kinases and Uses
/ TITLE OF INVENTION: Therefor
/ FILE REFERENCE: 35800/209996
/ CURRENT APPLICATION NUMBER: US/09/799,875
/ PRIOR FILING DATE: 2001-03-06
/ PRIOR APPLICATION NUMBER: 60/182,059
/ PRIOR FILING DATE: 2000-02-11
/ PRIOR APPLICATION NUMBER: 09/659,287
/ PRIOR FILING DATE: 2000-09-12
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 1074
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-799-875-9

Query Match 52.0%; Score 1070.8; DB 9; Length 1074;

Best Local Similarity 99.8%; Pred. No. 1,46-221;
Matches 1072; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 49 ATGCGAGCCACCCCTCTGTGCTGTCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 108
Db 1 ATGCGAGCCACCCCTCTGTGCTGTCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 60
Qy 109 TTTGATGACAACTTAAGTACCAGAGCTCCGCTCCAGAAAGAGCTCCAGTGGGCTCCAG 168
Db 61 TTTGATGACAACTTAAGTACCAGAGCTCCGCTCCAGAAAGAGCTCCAGTGGGCTCCAG 120
Qy 169 CCCAGACTGCCCCCTCTGTGCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 228
Db 121 CCCAGACTGCCCCCTCTGTGCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 180
Qy 229 GCTGTGCTCACTGCTCTCCCTCTTGTGGCTCTATGTCTCTGTGAGCCGAGAGCGGCTG 288
Db 181 GCTGTGCTCACTGCTCTCCCTCTTGTGGCTCTATGTCTCTGTGAGCCGAGAGCGGCTG 240
Qy 289 CCGGCTTAACCGGCTCTGTGCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 348
Db 241 CCGGCTTAACCGGCTCTGTGCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 300
Qy 349 GTCCAGAGAGCCCTGTGCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 408
Db 301 GTCCAGAGAGCCCTGTGCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 360
Qy 409 GCTGTGCTCACTGCTCTCCCTCTTGTGGCTCTATGTCTCTGTGAGCCGAGAGCGGCTG 468
Db 361 GCTGTGCTCACTGCTCTCCCTCTTGTGGCTCTATGTCTCTGTGAGCCGAGAGCGGCTG 420
Qy 469 CATGCGGACATGACCAAGCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 528
Db 421 CATGCGGACATGACCAAGCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 480
Qy 529 GTGTCTTTCGCGCAGATGAGCAACCGCTGTGCGCACTGTGCAAGAGAGCGGTTGAG 588
Db 481 GTGTCTTTCGCGCAGATGAGCAACCGCTGTGCGCACTGTGCAAGAGAGCGGTTGAG 540
Qy 589 GGTGTGTCTGTGGCCCACTGGAAGTCCCAAGTGGGACTTTCTGTGGGACACTTGGGCT 648
Db 541 GGTGTGTCTGTGGCCCACTGGAAGTCCCAAGTGGGACTTTCTGTGGGACACTTGGGCT 600
Qy 649 GAGAAGCTGAGAGCTCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 708
Db 601 GAGAAGCTGAGAGCTCTGTGCTGTGCGGTTTCTGTGCAAGAGAGCGGTTGAG 660
Qy 709 GCGTGTCCCAAGCTTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 768
Db 661 GCGTGTCCCAAGCTTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 720
Qy 769 GCAAGCCGATGTCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 828
Db 721 GCAAGCCGATGTCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 780
Qy 829 TTTCAAGACTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 888
Db 781 TTTCAAGACTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 840
Qy 889 CTTGAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 948
Db 841 CTTGAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 900
Qy 949 GCTGAAACGCTCAACGACCAAGGATCTTCTGTGAGCTGTGAGCTGTGAGCTGTGAG 1008
Db 901 GCTGAAACGCTCAACGACCAAGGATCTTCTGTGAGCTGTGAGCTGTGAGCTGTGAG 960
Qy 1009 CCGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 1068
Db 961 CCGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 1020
Qy 1069 GCGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 1122


```
RESULT 9
US-09-873-367C-341/C
; Sequence 341, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endrease, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 341
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-341

Query Match      15.3%; Score 315.4; DB 13; Length 353;
Best Local Similarity 95.2%; Pred. No. 1.7e-58;
Matches 336; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY      1644 TGGGAAGCTGTGTT-CCGAGCATCTGTCCTGTTGATTAAAGATTCCCTCCAGGCT 1702
DB      353 TGGGAAGCTGTGTTCCCGAGCATCTGTCCTGTTGATTAAAGATTCCCTCCAGGCT 294
QY      1703 AAGCTGGGATTGGGCGCAGAGATAAATCCAACTATAGAGGCTAGTCTTGCTTAAT 1762
DB      293 AAGCTGGGATTGGGCGCAGAGATAAATCCAACTATAGAGGCTAGTCTTGCTTAAT 234
QY      1763 CAAAGCTGTTCTGAAATGAGGCTCCAGGCTGTCAACCATYGGGGCTTCTGACCTGAGCAG 1822
DB      233 CAAAGCTGTTCTGAAATGAGGCGCCAGGCTGTCAACCATYGGGGCTTCTGACCTGAGCAG 174
QY      1823 CAAAGTTAGGAGCAGAGATTAGGCGAGGCTGTGCTGTCGCAACCTGGGAAAGTCCCAAGT 1882
DB      173 CAAAGTTAGGAGCAGAGATTAGGCGAGGCTGTGCTGTCGCAACCTGGGAAAGTCCCAAGT 114
QY      1883 GGGACTCTTCTGGGAGCACTTGGGGTCCCAATCCCAAGGCTCATACTCTAGGTTTGGAT 1942
DB      113 GGGACTCTTCTGGGAGCACTTGGGGTCCCAATCCCAAGGCTCATACTCTAGGTTTGGAT 54
QY      1943 ACCATGAGTATGATGTTTACTCTGTGCTTAATTAAGAGAAATTAAGAAATAA 1995
DB      53 ACCATGAGTATGATGTTTACTCTGTGCTTAATTAAGAGAAATTAAGAAATAA 1

RESULT 10
US-09-925-301-525
; Sequence 525, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
```

```
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 525
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (515)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (526)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (557)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-525
```

```
Query Match      13.1%; Score 268.8; DB 9; Length 562;
Best Local Similarity 96.7%; Pred. No. 2.5e-48;
Matches 267; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 GCTTGAGCCCCCGCGCGCGCCCGGCGCCACCGGAGAACGAGCGGCGAGATGCGAGCCACC 60
DB      274 GCTTGAGCCCCCGCGCGCGCCCGGCGCCACCGGAGAACGAGCGGCGAGATGCGAGCCACC 333
QY      61 CCTGCGCTGCTCCCGCGGGGTTCCCTGTCAGGAAGCGGTTGAGTTGATGATACAC 120
DB      334 CCTGCGCTGCTCCCGCGGGTTCCTGTCAGGAAGCGGTTGAGTTGATGATACAC 393
QY      121 TTAGATACGAGCGCTCCCGTCAGAAACGAGCTCGAAGTGCGGCCCGCCAGCCAGACTGCCC 180
DB      394 TTAGATACGAGCGCTCCCGTCAGAAACGAGCTCGAAGTGCGGCCCGCCAGCCAGACTGCCC 453
QY      181 CCTGCGCTGTTCCCGTCAGGCCCACTACTGTCAGATTCGTGCAACTGCTGTGGCACT 240
DB      454 CCTGCGCTGTTCCCGTCAGGCCCACTACTGTCAGATTCGTGCAACTGCTGTGGCACT 513
QY      241 GCTCCGCTTGGGCGCCCTATGTCCTCTGAGAGCCC 276
DB      514 GCTCCGCTTGGGCGCCCTATGTCCTCTGAGAGCCC 549
```

```
RESULT 11
US-10-291-172-116
; Sequence 116, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21372-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 116
; LENGTH: 1909
; TYPE: DNA
```


[illegible]

RESULT 13

```

US-10-084-817-300
Sequence 300, Application US/10084817
Publication No. US20030119009A1
GENERAL INFORMATION:
APPLICANT: Susan Stuart
APPLICANT: Jed G. Nuchtern
APPLICANT: Sharon B. Plon
APPLICANT: Jason M. Shohet
TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
FILE REFERENCE: PA-0046 US
CURRENT APPLICATION NUMBER: US/10/084,817
CURRENT FILING DATE: 2002-02-25
PRIOR APPLICATION NUMBER: 60/770,784
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 365
SOFTWARE: PERL Program
SEQ ID NO 300
LENGTH: 4336
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030119009A1 234427.4
US-10-084-817-300

```

Query Match	12.3%	Score 253	DB 15	Length 4336
Best Local Similarity	55.7%	Pred No	1.5e-44	
Matches 484	Conservative	0	Mismatches 385	Indels 0
			Gaps	0

Oy	135	CTTGAGGCCA	CTTACTGCTCCAGATCGTGCACCTGCTGGCCACTGCTCCCGCTTGG	254
Db	1459	CCTCGGCTCTCCCGAACCCGCCGAGACTCCGAATCTTGTCGACTTGCCTTCTTGATACGG	1518	
Oy	255	GCCCTATGTCCCTCTGAGAGCCCGAGAGAGGGGGGGGGCGTACCGGGGCTGCACTGGCC	314	
Db	1519	GAATACTTATTGTTGGAACTCTGAGAGGAGACCACTTTTTCGTGCGGTCATCGCA	1578	
Oy	315	TACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGAGAGCCGTGGCCGTGCA	374	
Db	1579	CAGCGAGAGAGAGCTGTGTGTGCAAGGTGTGATATCAGCTGTACAGGAATCCCTGGC	1638	
Oy	375	GCCCTACGGCGGGCGTGGCCCGGCAAGAGATGTGGCTGGCCCATGAGTCTGTGGCTGG	434	
Db	1639	ACCGGTCTTTTGGCTGTCTGCTCATAGTAACTCAACCAATCACTGAAATATATCTGGG	1698	
Oy	435	TACCCAGCTCTCTACGCGCTTTTTCACCTCGGACCCATGGGGACATGCAAGGCTGTGGCG	494	
Db	1699	TGAGACCAAAAGCTATGTGTCTTGAGCGAAAGCTATGGGGACATGTACATTCCTTGCTCG	1758	
Oy	495	AAGCCGCCACCGTATCCCTGAGCCCTGAGAGGCTGGCGTCTCTCCGCCAGATGGCCACCGC	554	
Db	1759	CACCTGCAGAAAGCTGAGAGAGAGAGAGGACCGCCAGACTGTTCATCCAGATTTGCTCTGGC	1818	
Oy	555	CCTGGCGCACTGTCAACAGCAACGATCTGGTCTCTGCTGATCTCAAGCTGTGTGCTTTGT	614	
Db	1819	AGTGGCCACTGGCCATGACGGGGGGCTGGTGTCTGGGACCTCAAGCTGGCGGAATTCAT	1878	
Oy	615	CTTGCGCTACCGTGAAGAGAAAGACTGTGTCTGAGAACTGTGAGAACTCTGTGGTGTCT	674	
Db	1879	CTTTAAGGACGAAAGAGAGACTCGGGTCAAGCTTGAAAGCCCTGGAAGACGCTTACATCTT	1938	

Oy	675	GACTGGGGCAGATGATATTCCTCGTGGGACAGACAGCGTGGCCAGAGCCTTACGTGGGACCTGA	734
Dd	1939	GCGGGAGAGATGATATATTCCTCTTCGACAAAGCATGCTGGCCGGCTTACGTAAAGCCAGA	1998
Oy	735	GATACTGAGCTCAGCGGGCCTCATATCTCGGGCAAGGACCGGATGTCTGGAGCCTGGGCGT	794
Dd	1999	GATCTTGAACACCAAGTGGGAGCTACTCGGGCAAAAGAGCCGAGCGTGTGGAGCCTGGGGGT	2058
Oy	795	GGCGCTCTTACCATGCTGGCGCGGCACATACCCCTTCCAGAGACTCGGAGCTGTCTGTCT	854
Dd	2059	GATGCTGTAAACCAATGTTGGTGGGGCGGTACCCCTTTCATAGACATTTAAACCCAGCTCCCT	2118
Oy	855	CTTCGGGAGAATCCGGCGGGGGGCGCTAAGCGCTTGGCTGACAGGCGCTTCGGGCGCTGGCCG	914
Dd	2119	CTTAGCAGAGATTCGGGCGGTGCCAGTTCAACATTTCCAGAGACTGTGTGCCCAAGGCCCA	2178
Oy	915	CTGTCTGTATTCGCTGCTCCCTCTTCGTCCGGAGCCAGCTGAACCGCTCAAGCCACAGGCAT	974
Dd	2179	GTCGCTCATCGAAGAGATTTCTGGGTGGGAGCCCTCAGACGGGTGAACCTCGACGAAAT	2238
Oy	975	CTCTCTGACACCCCTGCTGGGACGACAGAACCCGATGCCCTTAAAGCCCAACCCGATTCATCT	1034
Dd	2239	TCTGAGACATTCCTGTGGTTTCTACAGATTTTAAAGCGTTCGAAATTCAGCATATGTGTCTAA	2298
Oy	1035	CTGGGAGGCTGCGCCAGGTGGCTCCCTGATG	1063
Dd	2299	GGAAGTGTCTGACCAAGCTGGTGGCCGACG	2327

RESULT 14

```

US-10-044-090-269
; Sequence 269, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044, 090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 269
; LENGTH: 3280
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1820904CB1
US-10-044-090-269

```

Query Match	12.2%	Score 251.4	DB 14	Length 3280
Best Local Similarly	55.6%	Pred. No. 3e-44		
Matches 483, Conservative	0	Mismatches 386	Indels 0	Gaps 0

OY	195	CCTGAGCCCACTTACTGTCCAGATCGTGCAACTGCTGTGCACATGCTCCCGTCTTG	254
Db	403	CTCTGGCTCCCCGAGCCCGCCGAGACTCCGAACTTGTGGATTGCTTTCTTGTATCGG	462
OY	255	GCCCTATGTCTCTCTGAGCCCGAGAGAGGGGGGGCCGACCTACCGGGGCTCTGCATGGCC	314
Db	463	GAATATCTTATTATTGTGGAACTCTGGAGGGAGACCAAGTTTGTGTGCGTGTGATCTGCA	522
OY	315	TACAGGCACGTAGATATACCTGCAGAGGTACCCCGTCCAGGAACCCCTGGCCGTGTGGA	374
Db	523	CAGCGGAGAGGAGCTGTGTGTGCAGAGGTGTTGATATCAGCTGTACAGAGAACTCCGTGGC	582
OY	375	GCCTTACGCGGGCTGCCCCCGGCACAAGCATGTGAGCTGTGGCCCATGTAGGTCTGGCTGG	434
Db	583	ACCGTGTCTTTTGCCTGTCTGTGTCTATTAATCATCAACCAATTCATGAAATTAATCTGGG	642
OY	435	TACCACTCTCTACAGCTTTTCACTGTGAGCCATGTGGGACATGCAAGACTGTGTGG	494

Db 643 TGAGACCAAGCCTATGTTCTTTGAGCGAAGCATGAGCATTCCTTCGCG 702
Qy 495 AAGCGCCACCGTATCCCTGAGCCTGAGCGCTGCTCTTCCGCGAGATGCGCAGCC 554
Db 703 CACCTGCAAGAGCTGAG 762
Qy 555 CTTGCGGACCTGTCACGAG 614
Db 763 AGTGGCCCTGCTGAG 822
Qy 615 CTTGCGGACCTGTCACGAG 674
Db 823 CTTTAAAG 882
Qy 675 GACTGGGACAGATATTCCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734
Db 883 GCGGGAGAGATGATGATTCCTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 942
Qy 735 GATACGAGCTCAAG 794
Db 943 GATCTTGAACAG 1002
Qy 795 GCGCCTCTTCAAG 854
Db 1003 GATCTGTAACAG 1062
Qy 855 CTTGCGGAGATCCG 914
Db 1063 CTTGAG 1122
Qy 915 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 974
Db 1123 GTGCTTATCCGAG 1182
Qy 975 CCTCTGACAG 1034
Db 1183 TCTGAGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1242
Qy 1035 CTGGAG 1063
Db 1243 GGAAGTGTGAG 1271

RESULT 15
US-10-240-965-168
Sequence 168, Application US/10240965
Publication No. US20030165924A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: SHIRPMAN, DOV
APPLICANT: SOMOGYI, Roland
APPLICANT: LAWN, Richard M.
APPLICANT: SEILHAMER, Jeffrey J.
APPLICANT: PORTER, Gordon J.
APPLICANT: MIKITA, Thomas
APPLICANT: TAI, Julie
TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
FILE REFERENCE: PA-0025 PCT
CURRENT APPLICATION NUMBER: US/10/240,965
CURRENT FILING DATE: 2002-10-04
PRIORITY APPLICATION NUMBER: 60/195,106
PRIORITY FILING DATE: 2000-04-05
NUMBER OF SEQ ID NOS: 276
SOFTWARE: PERL Program
SEQ ID NO 168
LENGTH: 3324
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: mlec_feature
OTHER INFORMATION: incyle ID No. US20030165924A1 197301.4
NAME/KEY: unsure

LOCATION: 254, 378
OTHER INFORMATION: a, t, c, g, or other
US-10-240-965-168
Query Match 12.0%, Score 247, DB 13, Length 3324,
Best Local Similarity 61.1%, Pred. No. 2.7e-43,
Matches 400, Conservative 0, Mismatches 255, Indels 0, Gaps 0
Qy 338 AGGTGATCCCGTCCAG 397
Db 635 AGGTGATCCCGTCCAG 694
Qy 398 ACAAGCATGAG 457
Db 695 ACAGCAACATTAAG 754
Qy 458 TCACCTGAG 517
Db 755 TTGAG 814
Qy 518 CTGAG 577
Db 815 AGGAG 874
Qy 578 GTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637
Db 875 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 934
Qy 638 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 697
Db 935 AGCTTGAAG 994
Qy 698 GAG 757
Db 995 CAG 1054
Qy 758 ACTGGGAG 817
Db 1055 ACTCGGAG 1114
Qy 818 GCCACTACCCCTTCCAG 877
Db 1115 GACGATACCCCTTCCAG 1174
Qy 878 CTTAGGCTTGCCTGAG 937
Db 1175 AGTCTGATCTGAG 1234
Qy 938 GTGGAG 992
Db 1235 GACGGAG 1289

Search completed: January 16, 2004, 20:19:03
Job time : 717.201 secs


```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25890
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25890

Query Match
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TPLAAPAG 11
Db 306 TPLAAPAG 313

RESULT 3
US-09-252-991A-18605
; Sequence 18605, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18605
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18605

Query Match
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 339 GLGLDEAR 346
Db 78 GLGLDEAR 85

RESULT 4
US-09-252-991A-31777
; Sequence 31777, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31777
; LENGTH: 67
; TYPE: PRT
```

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31777

Query Match
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 177 GLVLRDL 183
Db 16 GLVLRDL 22

RESULT 5
US-09-328-352-5099
; Sequence 5099, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Brelton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5099
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5099

Query Match
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 209 LTGPDDS 215
Db 10 LTGPDDS 16

RESULT 6
US-09-252-991A-20965
; Sequence 20965, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20965
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20965

Query Match
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PAGSLSR 15
Db 30 PAGSLSR 36

RESULT 7
US-09-384-162-16
; Sequence 16, Application US/09384162
```

Patent No. 6376747
GENERAL INFORMATION:
APPLICANT: Xing, Ti
APPLICANT: Malik, Kamal
APPLICANT: Martin-Heller, Teresa
TITLE OF INVENTION: No. 6376747e1 Plant-Derived Map Kinase Kinase
FILE REFERENCE: 08-884280US
CURRENT FILING DATE: 1999-08-27
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 133
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-384-162-16

Query Match 2.0%; Score 7; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 DWSLGV 249
DB 69 DWSLGV 75

RESULT 8
US-09-252-991A-24993
Sequence 24993, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24993
LENGTH: 146
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24993

Query Match 2.0%; Score 7; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 LLPLSP 53
DB 63 LLPLSP 69

RESULT 9
US-09-252-991A-17117
Sequence 17117, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17117
LENGTH: 153
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17117

Query Match 2.0%; Score 7; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PLAPAG 11
DB 109 PLAPAG 115

RESULT 10
US-09-482-273-211
Sequence 211, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
CURRENT FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 211
LENGTH: 160
TYPE: PRT
ORGANISM: Homo sapiens
US-09-482-273-211

Query Match 2.0%; Score 7; DB 4; Length 160;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 PCLLPL 50
DB 3 PCLLPL 9

RESULT 11
US-08-469-537A-83
Sequence 83, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisongierre, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Ga11 M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 163 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-83

Query Match 2.0%; Score 7; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 243 DWSLGV 249
Db 90 DWSLGV 96

RESULT 12
US-09-252-991A-23665
Sequence 23665, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23665
LENGTH: 167
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23665

Query Match 2.0%; Score 7; DB 4; Length 167;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 POPRLPP 45
Db 112 POPRLPP 118

RESULT 13
US-08-469-537A-35
Sequence 35, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisongierre, et al.

TITLE OF INVENTION: EHK AND ROR TYROSINE
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Ga11 M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 190 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-35

Query Match 2.0%; Score 7; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 243 DWSLGV 249
Db 103 DWSLGV 109

RESULT 14
US-08-469-537A-36
Sequence 36, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisongierre, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23665
LENGTH: 167
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-08-469-537A-36

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempster, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 190 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-36

Query Match 2.0%; Score 7; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 243 DWSLGV 249
DB 103 DWSLGV 109

RESULT 15
US-08-444-083-8
Sequence 8, Application US/08444083
Patent No. 5571675
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
City: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,083
FILING DATE: 17-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286304
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy B.

REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 894P1D5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-444-083-8

Query Match 2.0%; Score 7; DB 1; Length 201;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 282 AGLSAPA 288
DB 69 AGLSAPA 75

Search completed: January 15, 2004, 15:07:22
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OW protein - protein search, using sw model

Run on: January 15, 2004, 14:56:21 Search time 21 Seconds
(without alignments)
721.299 Million cell updates/sec

Title: US-09-909-474d-2

Sequence: 1 MRATPLAAGSLSRKRL.....GIGLDEAREEGREVLYG 358

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents, AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*
7: /cgn2_6/ptodata/1/1aa/6E.COMB.pep:*
8: /cgn2_6/ptodata/1/1aa/6F.COMB.pep:*
9: /cgn2_6/ptodata/1/1aa/6G.COMB.pep:*
10: /cgn2_6/ptodata/1/1aa/6H.COMB.pep:*
11: /cgn2_6/ptodata/1/1aa/6I.COMB.pep:*
12: /cgn2_6/ptodata/1/1aa/6J.COMB.pep:*
13: /cgn2_6/ptodata/1/1aa/6K.COMB.pep:*
14: /cgn2_6/ptodata/1/1aa/6L.COMB.pep:*
15: /cgn2_6/ptodata/1/1aa/6M.COMB.pep:*
16: /cgn2_6/ptodata/1/1aa/6N.COMB.pep:*
17: /cgn2_6/ptodata/1/1aa/6O.COMB.pep:*
18: /cgn2_6/ptodata/1/1aa/6P.COMB.pep:*
19: /cgn2_6/ptodata/1/1aa/6Q.COMB.pep:*
20: /cgn2_6/ptodata/1/1aa/6R.COMB.pep:*
21: /cgn2_6/ptodata/1/1aa/6S.COMB.pep:*
22: /cgn2_6/ptodata/1/1aa/6T.COMB.pep:*
23: /cgn2_6/ptodata/1/1aa/6U.COMB.pep:*
24: /cgn2_6/ptodata/1/1aa/6V.COMB.pep:*
25: /cgn2_6/ptodata/1/1aa/6W.COMB.pep:*
26: /cgn2_6/ptodata/1/1aa/6X.COMB.pep:*
27: /cgn2_6/ptodata/1/1aa/6Y.COMB.pep:*
28: /cgn2_6/ptodata/1/1aa/6Z.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1675	88.6	360	US-09-509-902A-9	Sequence 9, Appl
2	290.5	15.4	607	US-08-878-989-15	Sequence 15, Appl
3	290.5	15.4	607	US-09-272-796-15	Sequence 15, Appl
4	274.5	14.5	633	US-08-557-006C-43	Sequence 43, Appl
5	274.5	14.5	668	US-09-930-181-2	Sequence 2, Appl
6	268.5	14.2	252	US-07-857-224B-26	Sequence 26, Appl
7	265.5	14.0	339	US-08-688-988-33	Sequence 33, Appl
8	263.5	13.9	264	US-07-857-224B-19	Sequence 19, Appl
9	263	13.9	257	US-07-857-224B-25	Sequence 25, Appl
10	260.5	13.8	552	US-08-557-006C-40	Sequence 40, Appl
11	256	13.5	354	US-08-688-988-29	Sequence 29, Appl
12	253	13.4	264	US-07-857-224B-18	Sequence 18, Appl
13	252.5	13.4	433	US-08-913-050A-7	Sequence 7, Appl
14	252.5	13.4	433	US-08-749-902-5	Sequence 5, Appl
15	248	13.1	359	US-08-688-988-32	Sequence 32, Appl
16	247.5	13.1	685	US-08-878-989-1	Sequence 1, Appl
17	247.5	13.1	685	US-09-136-282-2	Sequence 2, Appl
18	247.5	13.1	685	US-09-272-796-1	Sequence 1, Appl
19	247.5	13.1	685	US-09-505-744-2	Sequence 2, Appl
20	246.5	13.0	382	US-09-142-551A-2	Sequence 2, Appl
21	244.5	12.9	604	US-09-523-849-35	Sequence 35, Appl
22	243.5	12.9	556	US-09-800-960-4	Sequence 4, Appl
23	243.5	12.9	603	US-09-800-960-2	Sequence 2, Appl
24	243.5	12.9	603	US-09-930-181-17	Sequence 17, Appl
25	243	12.9	273	US-08-252-995D-10	Sequence 10, Appl
26	243	12.9	273	US-08-834-108-10	Sequence 10, Appl
27	243	12.9	416	US-08-252-995D-2	Sequence 2, Appl

28	243	12.9	416	2	US-08-834-108-2	Sequence 2, Appl
29	243	12.9	464	1	US-08-252-995D-6	Sequence 6, Appl
30	243	12.9	464	2	US-08-834-108-6	Sequence 6, Appl
31	243	12.9	925	2	US-08-252-995D-4	Sequence 4, Appl
32	243	12.9	925	2	US-08-834-108-4	Sequence 4, Appl
33	242.5	12.8	351	3	US-08-688-988-28	Sequence 28, Appl
34	242.5	12.8	351	3	US-08-688-988-31	Sequence 31, Appl
35	241	12.7	295	1	US-07-951-715A-23	Sequence 23, Appl
36	241	12.7	295	2	US-08-459-448A-23	Sequence 23, Appl
37	241	12.7	295	3	US-08-459-595A-23	Sequence 23, Appl
38	241	12.7	295	3	US-08-459-504B-23	Sequence 23, Appl
39	241	12.7	295	3	US-08-459-444-23	Sequence 23, Appl
40	241	12.7	295	4	US-09-547-422-23	Sequence 23, Appl
41	240	12.7	776	4	US-09-523-849-34	Sequence 34, Appl
42	239	12.6	345	3	US-09-101-146-1	Sequence 1, Appl
43	237	12.5	363	3	US-08-688-988-30	Sequence 30, Appl
44	235.5	12.5	463	1	US-07-951-715A-25	Sequence 25, Appl
45	235.5	12.5	463	2	US-08-459-448A-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-509-902A-9
Sequence 9, Application US/09509902A
Patent No. 6387676
GENERAL INFORMATION:
APPLICANT: Virca, Duke
APPLICANT: Bird, Timothy A.
APPLICANT: Anderson, Dirk M.
APPLICANT: Marken, John S.
TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
FILE REFERENCE: 2877-US
CURRENT APPLICATION NUMBER: US/09/509,902A
CURRENT FILING DATE: 1999-08-03
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 9
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
US-09-509-902A-9

Query Match 88.6%; Score 1675; DB 4; Length 360;
Best Local Similarity 99.7%; Pred. No. 9.9e-164;
Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRATPLAAGSLSRKRLLEDDNDTERPVQKARSGQPLPCLPLSPPTAPDRAT	60
DB	43	MRATPLAAGSLSRKRLLEDDNDTERPVQKARSGQPLPCLPLSPPTAPDRAT	102
QY	61	AVATASRLGPLYLLEPEEGRAVRAHCPGTGTCTCKYYPVOBALAVLEPYARLPKHIV	120
DB	103	AVATASRLGPLYLLEPEEGRAVRAHCPGTGTCTCKYYPVOBALAVLEPYARLPKHIV	162
QY	121	ARPTVLAGTQLVAFPTRTGDMHSLVRSRRIRTEPBAVTPROMATLAAHCHQGLVL	180
DB	163	ARPTVLAGTQLVAFPTRTGDMHSLVRSRRIRTEPBAVTPROMATLAAHCHQGLVL	222
QY	181	RDILKCRFVFPDRERKLVLENLEDSCVLTGPDSDLMKHAAPVAGPEIISRSYSYGK	240
DB	223	RDILKCRFVFPDRERKLVLENLEDSCVLTGPDSDLMKHAAPVAGPEIISRSYSYGK	282
QY	241	AADVSLGVALFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRRRP	300
DB	283	AADVSLGVALFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRRRP	342
QY	301	ABRLTAGIILHPWLROD 318	
DB	343	ABRLTAGIILHPWLROD 360	

RESULT 2
US-08-878-989-15
Sequence 15, Application US/08878989
Patent No. 5885803
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, Preeti
APPLICANT: Goli, Surya K.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSES: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTESQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,989
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 607 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1827450
US-08-878-989-15
Query Match 15.4%; Score 290.5; DB 2; Length 607;
Best Local Similarity 29.2%; Pred. No. 3.4e-21;
Matches 92; Conservative 45; Mismatches 129; Indels 49; Gaps 12;
QY 48 LPLSP---TAPDRATAVATASRLGPVYLLBPEBG--RAYRALHCPGTETCTCVYV 101
DB 5 LPTSDPGRITDPBSGRITLKGRLLG-----KGFARCYBATDTGTGSAAVAVIQQ 56
QY 102 QBALVLEPYARLP-----PHKVAPTEVLAQTOLLYAFTR--TGDHMSLVR 149
DB 57 SR---VAKPHQBEKILNIEILHRDLQHRHIVFSHHFEDADNIYIFLEICSRKSLAHYK 113
QY 150 SHRIPEPEAAVLFROMATALAHCHQGLVLDLKLCPVFAVDERKKVLLENLDSQVL 209
DB 114 AHTLLEPEVRYRYYRLGILSGLYKHQGIILHRDLKLGNFPII--EMMBLKVGDFGLAARL 171
QY 210 TGPDDSLMKHAC--PAYGPEILSSRASYSGRKAADVWSLGVALFTMLAGHYFPQDSSEV 267
DB 172 BPPEOR--KRTICGTBNVAVPEVLLRQG--HGPEADVWSLGCVMYTLTGSPPEFTADLK 227

QY 268 LIFGKIRGAVAPAGLSAPARCLVRCILRRBPARRLTATGILLHPWLKQ----DPMPLA 323
DB 228 EYRCIKQVHYVLPASLSLPARQLAAIIRKASPRDRPSIDQILRHDPFTKGTTPDRPLPS 287
QY 324 PTRSHLEMAQVVPD 338
DB 288 -----SCVTVPD 294

RESULT 3
US-09-272-796-15
Sequence 15, Application US/09272796
Patent No. 6207148
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, Preeti
APPLICANT: Goli, Surya K.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSES: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTESQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/272,796
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/878,989
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 607 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1827450
US-09-272-796-15
Query Match 15.4%; Score 290.5; DB 3; Length 607;
Best Local Similarity 29.2%; Pred. No. 3.4e-21;
Matches 92; Conservative 45; Mismatches 129; Indels 49; Gaps 12;
QY 48 LPLSP---TAPDRATAVATASRLGPVYLLBPEBG--RAYRALHCPGTETCTCVYV 101
DB 5 LPTSDPGRITDPBSGRITLKGRLLG-----KGFARCYBATDTGTGSAAVAVIQQ 56
QY 102 QBALVLEPYARLP-----PHKVAPTEVLAQTOLLYAFTR--TGDHMSLVR 149
DB 57 SR---VAKPHQBEKILNIEILHRDLQHRHIVFSHHFEDADNIYIFLEICSRKSLAHYK 113

QY 150 SRRRIPEBAVLPFROMATALAHCHQGLVLRDLKLCRFVADREKKLVLENLEDSCTL 209
DB 114 AHHTLLEPEVRYRYLGLKYLHQRGLHRLDKGNPPT--EMMELKVGDFGLAARL 171
QY 210 TGPDDSLMDKHAC--PAYVGPETILSSRASYSCKADAVWSLGVALLFTMLAGHYFQDSSEPV 267
DB 172 BPEBQR--KKTICGFNNYVAPEVLNQG--HGEADVWSLGCVMYTLGSPPEFTADLK 227
QY 268 LIFGKIRGAYVLPAGLSAPARCLVRCILRRPARELTATGILLHPMLRQ---DPMPLA 323
DB 228 ETRYRCIKQVHYTLPLASLIPAROLLALIRASPRDRPSIDQLRHDFFTKGYTPDRPLPIS 287
QY 324 PTRSHLMEAAOVVPD 338
DB 288 -----SCVTVPD 294

RESULT 4
US-08-557-006C-43
Sequence 43, Application US/08557006C
Patent No. 6258547
GENERAL INFORMATION:
APPLICANT: Berl, Rajindar K.
APPLICANT: Carling, David
APPLICANT: Forster, Robert A.
TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NGAP/PHM37586/US7
CURRENT APPLICATION NUMBER: US/08/557, 006C
PRIOR FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9318010.7
PRIOR FILING DATE: 1993-08-31
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43
LENGTH: 633
TYPE: PRT
ORGANISM: Yeast
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)-(633)
OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

Query Match 14.5%; Score 274.5; DB 3; Length 633;
Best Local Similarity 27.8%; Pred. No. 1.6e-19;
Matches 88; Conservative 53; Mismatches 138; Indels 37; Gaps 11;

QY 51 SEPTAPDATAVATASRLGPLYLLP-BEG--GRAYRALHCPGTGYTCVYPVOBALAV 107
DB 35 SNSTLANPKESLADGAHIGNVQIVKTLGEGSGFYGLAVHTTGKVALKIKIN--KVALAK 93
QY 108 LE-----PYALLPPIHKVAPRTVYLAGQLVAFPTTRHGMHSLVSRKHAIPPE 158
DB .94 SDMGRIEREISYLLLRHPHILIKYDVYKSKDEIIMVIEVAGNLSFDYIVQDRDKMSQE 153
QY 159 AAVLPROMATALAHCHQGLVLRDLKLCRFVADREKKLVLENLEDSCTLGPDSDSLMD 218
DB 154 ARRFQOIIISAVEYCHRKIKIVRDLPENLLDEHLNLIKIDPGL--SNMT---DGNPL 208
QY 219 KHAC--PAYVGPETILSSRASYSCKADAVWSLGVALLFTMLAGHYFQDSSEPVLLFGKIRG 276
DB 209 KTSGSPNYAAPKVISGKL--YAGPEVDWSCGVIIYVMLCRLLPDDSDSIPVLFKNISNG 267
QY 277 AYALPAGISAPABCLVRCLLRREPARLTATGILLHPMLRQD--PMPPLAPTSRLMEAAOV 335
DB 268 VYTLPEKLSPPAGGLIKMLLVNPLNRSIHEIMODDFKVDLPYLLP----- 316
QY 336 VPDGLGLDEARBEGD 351

DB 317 -PD---LKPHEENE 328

RESULT 5
US-09-930-181-2
Sequence 2, Application US/09930181
Patent No. 6455292
GENERAL INFORMATION:
APPLICANT: Origene Technologies
TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
FILE REFERENCE: 16U 101 V1
CURRENT APPLICATION NUMBER: US/09/930,181
PRIOR FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 668
TYPE: PRT
ORGANISM: Homo sapiens
US-09-930-181-2

Query Match 14.5%; Score 273.5; DB 4; Length 668;
Best Local Similarity 27.6%; Pred. No. 2.2e-19;
Matches 92; Conservative 52; Mismatches 138; Indels 51; Gaps 14;

QY 65 ASRLGPLYL--LEPEBGRAYRALHCPGTGYTCV-----YVGBALAVLP 110
DB 13 AQVGPYRLKLTGKGTGLVGLVGVHCVTCQKVAIKIVREKLSBSVLKMKVEREIAL-- 70
QY 111 VARLPPIHKVAPRTVYLAGQLVAFPTTR--THGDMHSLVSRHRIPEBAVLPFROMAT 169
DB 71 --KLIEHPVVLKLVHVENKTYLVLEHVSGBLFDYLVKGRLLPCKARKEFRQIISA 128
QY 170 LAHCHQGLVLRDLKLCRFVADREKKLVLENLEDSCTLGPDSDSLMDKHAC--PAYVG 227
DB 129 LDFCHSHSICHRDLKPENNLL--DEKNIRIADFGMASIQVG--DSLET--SCGSPHYAC 183
QY 228 PEILSSRASYSCKADAVWSLGVALLFTMLAGHYFQDSSEPVLLFGKIRGAYVLPAGLSAP 287
DB 184 PEVIRGE--KYDGRKADVWSGCVILPALLVGLALPDDDNLRQLLEKVRGVFPHNPHPIPD 242
QY 288 ARCLVRCILRRPARELTATGILLHPVL---RQDPMPLAP-----TRSHLMEAAOVVPD 338
DB 243 CQSLRGMIENVDAARLTLEHILQKHIWYIGKNIEPEQPIPRKVOIRS--LPSLEDIDPD 301
QY 339 -----GLG-----LDEAREEGDEREVLY 357
DB 302 VLDMSHSLGCFRDRKLLQDLLEENQEKMIY 334

RESULT 6
US-07-857-224B-26
Sequence 26, Application US/07857224B
Patent No. 5958784
GENERAL INFORMATION:
APPLICANT: Benner, Steven A.
TITLE OF INVENTION: Predicting Folded Structures of Proteins
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSER: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8092
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/857,224B

FILED DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2830
TELEFAX: (International) 41 1 262 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 252
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Schizosaccharomyces pombe
FEATURE: Protein kinase; Table 8 Column 29
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanke, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
US-07-857-224B-26

Query Match 14.2%; Score 268.5; DB 2; Length 252;
Best Local Similarity 31.5%; Pred. No. 1.8e-19;

Matches 74; Conservative 44; Mismatches 108; Indels 9; Gaps 6;

QY 85 ALHCPGTGYTCVYVQVQALAVLE-PYARLPYHKNVAPTEVLATGQLLY-AFPTRTNG 142
DB 20 AKAKTGDLAIKIIPIRYASIGWEILMMLLHPNLRLYDWTQHLYALEVYVDG 79
QY 143 DMSLSVSRHRIPEPEAAVLFRMATLACHQHGVLRLKICRFVADREKRVLEN 202
DB 80 ELPHYIRKGPRLEREAHYLSQILDVAHCHFRFRHRLKLENTIIRVNEQO--IKI 136
QY 203 LEDSCVLGTGPDSDLMQKACPA--YVQPEILSSRAYSYSGAADVMSLGVALLFTMLAGHY 260
DB 137 ADRCGMATVERNDSCLENY-CGSLHYLAPEIVSHK-PYRGAPADVMSGVILYSLSKLP 194
QY 261 PDSEPVLLFGKIRRGAYALPAGISAPARCLVRCLEAREAEULATGILLHPWL 315
DB 195 FGQNTDVIYVKIRHGAVIDLPSISSAAQDLHRMLDVNPSTRITIPPEFSPHPL 249

RESULT 7

US-08-688-988-33
Sequence 33; Application US/08688988B
Patent No. 6096545
GENERAL INFORMATION:
APPLICANT: Lefebvre, Daniel D.
APPLICANT: Mabboudi, Mohammed A.
TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
FILE REFERENCE: PPL96-03
CURRENT APPLICATION NUMBER: US/08/688,988B
CURRENT FILING DATE: 1996-07-31
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PasteSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 339
TYPE: PRT
ORGANISM: Glycine max
US-08-688-988-33

Query Match 14.0%; Score 265.5; DB 3; Length 339;
Best Local Similarity 28.7%; Pred. No. 5.5e-19;
Matches 86; Conservative 35; Mismatches 86; Indels 93; Gaps 12;

QY 117 HKGVAPTEVLATGQLLYAFPTRTN-----GDMSLVSRHRIPEPEAAVLFRM 166
DB 58 HPIIRPKXV-----FLPTHLAIVLEYAAGSLERICNAGRLSDKARFPFOOL 108
QY 167 ATALACHQHGVLRLDKL-----CRFVADREKRVLENLSDSCVLGTG 211
DB 109 ISGVSYCHMOICHRCLKLENTLLDGNPAPRLKICDFGS-----KALLHSQKSYTGT- 163
QY 212 PDDSLMDKACPAYVGEPEILSSRAYSYSGAADVMSLGVALLFTMLAGHYPPDSEPVLLF- 270
DB 164 -----PAYIAPEVL-SRKEYDGKVDVMSGVITLYVWLGVAYPEDEDEDPNFR 211
QY 271 ----GKIRGAYALP-AGLSAPARCLVRCLEAREAEULATGILLHPMLRQD-PMPRL- 322
DB 212 KSIGRIMSVQYAIPIDVYVSKCERHLISRI FVANPARKIRINISEIKOHLFRKXLPREIIE 271
QY 323 -----APTRSHLWMAQVVPDGLL-DEAREEGDREV 354
DB 272 AERRGYBETQKQPSQSVBEIMQIIQEARTKIHGSA-----GTGTSVVRGDDEANBEV 326

RESULT 8

US-07-857-224B-19
Sequence 19; Application US/07857224B
Patent No. 5958784
GENERAL INFORMATION:
APPLICANT: Benner, Steven A.
TITLE OF INVENTION: Predicting Folded Structures of Proteins
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSER: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8092
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/857,224B
FILING DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2830
TELEFAX: (International) 41 1 262 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 264
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: rat
FEATURE: Protein kinase; Table 8 Column 21
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanke, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
US-07-857-224B-19

Query Match 13.9%; Score 263.5; DB 2; Length 264;

[illegible]

RESULT 9
US-07-857-224B-25

Sequence 25, Application US/07857224B
Patent No. 5958784

GENERAL INFORMATION:
APPLICANT: Bennett

TITLE OF INVENTION: Predicting Folded Structures of Proteins

! NUMBER OF SEQUENCES: 114
! CORRESPONDENCE ADDRESS:

ADDRESSEE: S
STREET: Had1

CITY: Zurich
STATE: none

COUNTRY: Switzerland

```

; ZIP: (NOTE: THIS IS A
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: Apple Macintosh

```

```

; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word

```

CURRENT APPLICATION DATA:

APPLICATION NUMBER
FILING DATE: 03/

CLASSIFICATION
PRIOR APPLICATION

TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2830

TELEPAX: (International)

! ELBA: none
! INFORMATION FOR SEQ ID NO: 25:

```
! SEQUENCE CHARACTERS 1-257
! LENGTH: 257
```

US-07-857-224B-25

Query Match 13.94; Score 263; DB 2; Length 257;

RESULT 10
INC-09-557

US-08-557-006C-40
; Sequence 40, Application US/08557006C

; PATENT NO. 6258547
; GENERAL INFORMATION:

; APPLICANT: Berl, Rajindar K.
; APPLICANT: Carling, David

APPLICANT: Forder, Robert A.
TITER OR INTENTION, NUCLEIC ACID ENCODING AND-ACTIVATED PROTEIN KINASE

FILE REFERENCE: NGAP/PHM37588/UST

; CURRENT APPLICATION NUMBER: US/08/557, 006
 ; CURRENT FILING DATE: 1996-03-06

; PRIOR APPLICATION NUMBER: PCT/GB94/0109
 ; PRIOR FILING DATE: 1994-05-20

PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-05-21

PRIOR APPLICATION NUMBER: G
PRIOR FILING DATE: 1993-03

1 PRIOR FILING
1 NUMBER OF SEQ

```

; SOFTWARE:
; SEQ ID NO

```

! LENGTH: 552
: TVDP. PPT

ORGANISM: Rat

```
! FEAT1URK:
! NAME/KEY: get
```

! LOCATION: (1)..(1747)
! OTHER INFORMATION: Full length cDNA sequence fragment of Human

Db 132 RHVVVVRDLKPEVVLDAQNNAKIADFGLSNMMSDGEPLRTSC-----GSPN 178
 Qy 225 YVGPETLSRASYSKGAADVMSLGVALLFTMLAGHYFPDSEPVLLGKIRGAYALPAGL 284
 Db 179 YAAPEVIGSRL-YAGEVDVIMSCGVLLYALLCCTLPFDEHVPFTLPKIRGVFIPEVL 237
 Qy 285 SAPARCLVRCILRREPARLTATGILLHPMLROD-PMPLAP 324
 Db 238 NRSIATLMMHLOVDPDKPATIKDIREHEWFKODLPSTLFP 278

RESULT 11

US-08-988-29
 / Sequence 29, Application US/08688988B
 / Patent No. 6096545
 / GENERAL INFORMATION:
 / APPLICANT: Lefebvre, Daniel D.
 / APPLICANT: Maibooobi, Mohammad A.
 / TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
 / FILE REFERENCE: PRL96-03
 / CURRENT APPLICATION NUMBER: US/08/688,988B
 / NUMBER OF SEQ ID NOS: 48
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 29
 / LENGTH: 354
 / TYPE: PRT
 / ORGANISM: Brassica napus
 US-08-688-988-29

Query Match 13.5%; Score 256; DB 3; Length 354;
 Best Local Similarity 29.5%; Pred. No.5.5e-18;

Matches 86; Conservative 32; Mismatches 104; Indels 70; Gaps 11;
 Qy 117 HKIVARPTB-VLAGTOLLYAFPTRTGDMHSLVSRHRIPEBAVLPROMATALAHCHQ 175
 Db 57 HPIIIFKEVVLPTFLALAMEYVAGGELPERICSGRSEDEBARYFQQLIGVGYCHA 116
 Qy 176 HGLVLDLKL-----CRFPADBERKKLVLENLEDCVLTGPPDSLMDKH 220
 Db 117 MOICHDLKLKENTLIDGSPAPRLKICDFGYS-----KSSLHSRPKSTVGT----- 162
 Qy 221 ACRAVGPETLSRASYSKGAADVMSLGVALLFTMLAGHYFPDSEPVLLF-----GKIRRG 276
 Db 163 --PAYIAPFVL-SRREYDGMADVSCGTLVYMLVGAFPEDEBPYKFRKTIQIMAV 219
 Qy 277 AYALP--AGLSAPARCLVRCILRREPARLTATGILLHPW-LRODPMPLAPTRSHLM--- 330
 Db 220 QYKIPYVHIISQCKHLISRIFVANSIKRITITAIKHPWFLKRLPRELLETQAAYFPK 279
 Qy 331 -----BAQVPPD-----GLGL-----DEAREEGDREVV 355
 Db 280 ENPTFSPQTAEBIKIIVDDAKTPPVVSRISIGFGMGKGDEEBEYDREVV 331

RESULT 12

US-07-857-224B-18
 / Sequence 18, Application US/07857224B
 / Patent No. 5958784
 / GENERAL INFORMATION:
 / APPLICANT: Benner, Steven A.
 / TITLE OF INVENTION: Predicting Folded Structures of Proteins
 / NUMBER OF SEQUENCES: 114
 / CORRESPONDENCE ADDRESS:
 / ADDRESSES: Steven A. Benner
 / STREET: Hadlaubstrasse 151
 / CITY: Zurich
 / STATE: none
 / COUNTRY: Switzerland
 / ZIP: (note: this is an international post code) CH-8092
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
 / COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 7.0
 SOFTWARE: Microsoft Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/857,224B
 FILING DATE: 03/25/92
 CLASSIFICATION: 436
 PRIOR APPLICATION DATA: none
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (international) 41 1 632 2830
 TELEFAX: (international) 41 1 262 2437
 TRILX: none
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 264
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE:
 DESCRIPTION: protein
 ORIGINAL SOURCE:
 ORGANISM: rat
 FEATURE: Protein kinase; Table 8 Column 20
 PUBLICATION INFORMATION:
 AUTHORS:
 AUTHORS: Hanks, S. K.
 AUTHORS: Quinn, A. M.
 TITL: The protein kinase family
 JOURNAL: Science
 VOLUME: 241
 PAGES: 42-52
 DATE: 1988

US-07-857-224B-18

Query Match 13.4%; Score 253; DB 2; Length 264;
 Best Local Similarity 29.4%; Pred. No.7.3e-18;

Matches 77; Conservative 49; Mismatches 116; Indels 20; Gaps 8;
 Qy 71 YLLPBBEGRAY-----RALHCPTEYTCRYPOBALA-----VLEPYA---RLPPHGH 119
 Db 3 YOLF-BLKGAFSVVRKCVLAGOEYAAKINTKKLSARDHOKLREARICRLKHNP 61
 Qy 120 VAPTEVLAGTOLLYAFPTR-THGDMHSLVSRHRIPEBAVLPROMATALAHCHQGL 178
 Db 62 IYRLDISISBGNHILIPLVYTGELFEDIVARBYSSADASHCIIQILBAYLHKOMGV 121
 Qy 179 VLRLDKLGRFVPADBERKKLV-LENLEDCVLTGPPDSLMDKHACPAVYGPETLSRASYS 237
 Db 122 VARDLKPETLLASLKGAAVTLADFGLAIBVEGEQAWFGPAGTGYLSPVL--RKOP 179
 Qy 238 SGKADVMSLGVALLFTMLAGHYFPDSEPVLLGKIRGAYALP-----GLSAPARCLVR 293
 Db 180 YKQPVDLMAAGVILYLLVGYPPFWDEDOHRLYQOI KGAADYFPSPBMDTVTPBAADLIN 239
 Qy 294 CLLRREPARLTATGILLHPML 315
 Db 240 KMLTINPSKRTIAPALAKHPWI 261

RESULT 13

US-08-913-050A-7
 / Sequence 7, Application US/08913050A
 / Patent No. 5837726
 / GENERAL INFORMATION:
 / APPLICANT: NESZU, Jun-ichi
 / TITLE OF INVENTION: DNA ENCODING PROTEIN KINASE
 / NUMBER OF SEQUENCES: 10
 / CORRESPONDENCE ADDRESS:
 / ADDRESSES: BROWDY AND NEIMARK, P.L.L.C.
 / STREET: 419 7th Street N.W., Suite 300
 / CITY: Washington
 / STATE: D.C.
 / COUNTRY: USA
 / ZIP: 20004

[illegible]

DB	2040	GAAATTGAGAAATATAA	2099
QY	2041	AAAAAAAAAAAAAAAAAAAAA	2059
Db	2100	AAAAAAAAAAAAAAAAAAAAA	2118
RESULT 2			
LOCUS	BC027484	2283 bp	mRNA linear PRI 08-APR-2002
DEFINITION	Homo sapiens, chromosome 20 open reading frame 97, clone MGC:34909		
ACCESSION	IMAGE:5104452, mRNA, complete cds.		
VERSION	BC027484		
KEYWORDS	MGC.		
SOURCE	BC027484.1 GI:20071610		
ORGANISM	Homo sapiens (human)		
REFERENCE	Bukacinski, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 2283)		
JOURNAL	Stranberg, R.		
REMARK	Direct Submission		
COMMENT	GenBank (04-APR-2002) National Institutes of Health, Mammalian Genome Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
	NIH-MGC Project URL: http://mgc.nci.nih.gov		
	Contact: MGC help desk		
	Email: gcgabs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center		
	Center code: BCM-HGSC		
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/		
	Contact: amg@bcm.tmc.edu		
	Gunnarctine, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.		
FEATURES			
source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Series: IRAK Plate: 52 Row: F Column: 11.		
	Location/Qualifiers		
	1..2283		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="locusID:57761"		
	/db_xref="taxon:9606"		
	/clone="MGC:34909 IMAGE:5104452"		
	/issue type="Cervix, Carcinoma"		
	/clone_id="NH MGC_12"		
	/lab host="DH10B"		
	/note="Vector: pCMV-SPORT6"		
CDS	295..1377		
	/codon_start=1		
	/product="Chromosome 20 open reading frame 97"		
	/protein_id="AA27484.1"		
	/db_xref="GI:20071611"		
	/translation="MRATPLAAPAGSLSRKKRLDNDLTBRPVOKRASGPOLP PCLLSPRPAPDRATAVATASRLGYULLEPRSGRAYRALHCPTGTCTCYTYVO BALVAIEFPAKLEPPKHGVARFTLVLTGLTQLIAVFTRTHODMSLVRSRRITPEPAA VLFRQWATALAHCHQGILVARDKCRFPVPADERRKCLVLENLSDSCVLTGPSDSLMD KACLPAYVGPEIISPSASYSGKAADWSLGVALFTMLAGHYPPQDSBPVTLFKIRRG AYALPGALSFAACLIVRCILREBPARRLTATGILLIPMLRDPMPLAPTRSHLMBAAQ VVDDGGDLBARBEEDREVLYNG"		
BASE COUNT	463 a 702 c 656 g 462 t		
ORIGIN			
Query Match	98.6%	Score 2030.6,	DB 9, Length 2283,
Best Local Similarity	99.8%	Pred. No. 0;	
Matches 2033; Conservative	0;	Mismatches	4; Indels 0; Gaps 0;

Qy	1311	TCCTCACAAGATGACAAATCGGACTCCCTTGAGCGTGAACAACCTTTCCATGACATATAG	1441
Db	1627	TGCTCACAAGATGACAAATCGGACTCCCTTGAGCGTGAACAACCTTTCCATGACATATAG	1686
Qy	1441	TCACGTGTCAACACTGGGTACACTTGTATCAAGTGTGGGCTTCCATGTATGTGTGCTCA	1500
Db	1687	TCACGTGTCTAACACTGGGTACACTTGTATCAAGTGTGGGCTTCCATGTATGTGTGCTCA	1746
Qy	1501	GGCACCCTGTGTCCAAAGACATCCCTTTTCACAACAAACAGCTGCTTTGTATTTGTA	1566
Db	1747	GGCACCCTGTGTCCAAAGACATCCCTTTTCACAACAAACAGCTGCTTTGTATTTGTA	1806
Qy	1561	CGTTTTCAGAGAAAGGAGGTATCCCTGTGCAAAAGGCTCCAGGCTCTCCCTGTCAACT	1620
Db	1807	CCTTTTCAGAGAAAGGAGGTATCCCTGTGCAAAAGGCTCCAGGCTCTCCCTGTCAACT	1866
Qy	1621	CAGAACCCAAAGCCAGCTCACTCTTGGAACCTGTGTTCCAGCATCTGTGCTCTTGATT	1680
Db	1867	CAGAACCCAAAGCCAGCTCACTCTTGGAACCTGTGTTCCAGCATCTGTGCTCTTGATT	1928
Qy	1661	AAGAGATTCTCCTTCAGGCGCTTAAGCCTGGGATTTGGGCGAGAGATTAAGATCCAAACTA	1740
Db	1927	AAGAGATTCTCCTTCAGGCGCTTAAGCCTGGGATTTGGGCGAGAGATTAAGATCCAAACTA	1986
Qy	1741	TGAGGCTAGTCTTGTCTTAACTCAAGACTGTTCTGGAATGAGGCTCAGGCGCTGTCAAC	1800
Db	1987	TGAGGCTAGTCTTGTCTTAACTCAAGACTGTTCTGGAATGAGGCTCAGGCGCTGTCAAC	2046
Qy	1801	ATGGGGCTTTCTGACCTGAGACACAAAGTTGAGGAGCAGAGATTAAGCAGAGCTGTCTGT	1866
Db	2047	ATGGGGCTTTCTGACCTGAGACACAAAGTTGAGGAGCAGAGATTAAGCAGAGCTGTCTGT	2106
Qy	1861	GGCCACACTGGAAGGTCCAGGCTGGAGCTCTTCTGGGAGCACATTTGGGCTCCAAATCCAG	1920
Db	2107	GGCCACACTGGAAGGTCCAGGCTGGAGCTCTTCTGGGAGCACATTTGGGCTCCAAATCCAG	2166
Qy	1921	GTCCACTACTTGAAGTTTGGATACCATGATATGTATGTTTAACTGTGCTTATTAAGGA	1980
Db	2167	GTCCACTACTTGAAGTTTGGATACCATGATATGTATGTTTAACTGTGCTTATTAAGGA	2226
Qy	1981	GAAATATGAAT	2011
Db	2227	GAAATATGAAT	2257
RESULT 4			
AX099934		2116 bp	DNA
LOCUS	Sequence 16 from Patenc	W00120004.	
DEFINITION	AX099934		
ACCESSION	AX099934.1	GI:13538944	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1	Yue, H., Tang, Y.T., Bandman, O., Hillman, J.L., Baughn, M.R.,	
AUTHORS	Azizmay, Y. and Lu, D.A.		
TITLE	Protein phosphatase and kinase proteins		
JOURNAL	Patent: WO 0120004-A 16 22-MAR-2001;		
FEATURES	Incyte Genomics, Inc. (US)		
	Location/Qualifiers		
Source	1..2116		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID NO: 1271505CB1"		
BASE COUNT	407 a	663 c	609 g
ORIGIN		437 t	
Query Match	97.1%;	Score 1999.2;	DB 6, Length 2116;
Best Local Similarity	99.9%;	Pred. No. 0;	

	Matches	2001	Conservative	0	Mismatches	3	Indels	0	Gaps	0
QY	1	GCTTGAGCCCCGCGCGCCCGGAGCCCA	CGCGGAA	CGA	CGGCGGAGATCGAGCCACC	60				
DB	113	GCTCTAGAGCCCCGCGCGCCCGGAGCCCA	CGCGGAA	CGA	CGGCGGAGATCGAGCCACC	172				
QY	61	CCTCGGCTGTGCTCGTGGGGGTTCCTCTG	CCAGAA	GAA	GGGTTTGAGTTTGAGTGAACAAC	120				
DB	173	CCTCGGCTGTGCTCGTGGGGGTTCCTCTG	CCAGAA	GAA	GGGTTTGAGTTTGAGTGAACAAC	232				
QY	121	TTAGATACGAGCGTCCCGTCCAGAAA	CGAGCTCGAAGTGGGCC	CCAG	CCGACATGCGCC	180				
DB	223	TTAGATACGAGCGTCCCGTCCAGAAA	CGAGCTCGAAGTGGGCC	CCAG	CCGACATGCGCC	292				
QY	181	CCCTGCGCTGTGGCCCTCGAGGCCCA	CTA	CTG	CTCAGATTCGTGGAATCTGCTGTGGCCACT	240				
DB	293	CCCTGCGCTGTGGCCCTCGAGGCCCA	CTA	CTG	CTCAGATTCGTGGAATCTGCTGTGGCCACT	352				
QY	241	GCCCTCCGCTCTTGAGGACCTATGTCT	CTCTGAG	CCCGAGAGAGGCGGCGGAGCCCTA	CCG	300				
DB	353	GCCCTCCGCTCTTGAGGACCTATGTCT	CTCTGAG	CCCGAGAGAGGCGGCGGAGCCCTA	CCAG	412				
QY	301	GCCCTGCACTGCCCCCTACAGGCA	CTGAGTATAC	CTGCAAGGTGTAC	CCCCGTCCAGAGAAC	360				
DB	413	GCCCTGCACTGCCCCCTACAGGCA	CTGAGTATAC	CTGCAAGGTGTAC	CCCCGTCCAGAGAAC	472				
QY	361	CTGGCCGCTGTGAGAGCCCTACGAGCG	GCGTGC	CCCCCGGCAAGCAATGTGTGCTCGG	CCCACT	420				
DB	473	CTGGCCGCTGTGAGAGCCCTACGAGCG	GCGTGC	CCCCCGGCAAGCAATGTGTGCTCGG	CCCACT	532				
QY	421	GAGGTCTGTGCTGTGATACCCAGCT	CCTCTA	CGCCCTTTTTCAC	TCGGA	CCCATGTGGAGCATG	480			
DB	533	GAGGTCTGTGCTGTGATACCCAGCT	CCTCTCTA	CGCCCTTTTTCAC	TCGGA	CCCATGTGGAGCATG	592			
QY	481	CACAGCCCTGTGTGTGCGAAGCCGCA	CCGATATCCCTTGAGGCTGTG	CGCTGCTCTTCCGC		540				
DB	593	CACAGCCCTGTGTGTGCGAAGCCGCA	CCGATATCCCTTGAGGCTGTG	CGCTGCTCTTCCGC		652				
QY	541	CAGATGTGCAACCGGCCCTGTGCGCA	CTGTCTA	CCAGACA	CGGTCTGTGCTGTGATCTCAAG	600				
DB	653	CAGATGTGCAACCGGCCCTGTGCGCA	CTGTCTA	CCAGACA	CGGTCTGTGCTGTGATCTCAAG	712				
QY	601	CTGTGTGCTTTGTGTCTTGTGCTGAC	CCGTGAGAGAA	GAA	GCTGTGTCTGTGAGAACCTGTGAG	660				
DB	713	CTGTGTGCTTTGTGTCTTGTGCTGAC	CCGTGAGAGAA	GAA	GCTGTGTCTGTGAGAACCTGTGAG	772				
QY	661	GACTCTGCGCGCTGACATGGGCGAGAT	GATTC	CCCTGTGTGGGCAAGAC	ACGCGTCC	832				
DB	773	GACTCTGCGCGCTGACATGGGCGAGAT	GATTC	CCCTGTGTGGGCAAGAC	ACGCGTCC	832				
QY	721	TACGTGTGAGACCTGAGATATCTGAG	CTCA	CGGCGCTCATATCTG	GGGCAAGGCAAGCGCATGTCT	780				
DB	833	TACGTGTGAGACCTGAGATATCTGAG	CTCA	CGGCGCTCATATCTG	GGGCAAGGCAAGCGCATGTCT	892				
QY	781	TGAGAGCTGTGGGCGTGTGGGGCTCTT	CAACAATGTGTGGCGCGGCA	CTAC	CCCTTCCAGAGATCTG	840				
DB	893	TGAGAGCTGTGGGCGTGTGGGGCTCTT	CAACAATGTGTGGCGCGGCA	CTAC	CCCTTCCAGAGATCTG	952				
QY	841	GAGCGCTGTCTGTGCTCTTGTGCGCA	AGATCGCG	CGCGGAGGCTTAC	CGACTTGTGCTGTGCAAGGCTCTC	900				
DB	953	GAGCGCTGTCTGTGCTCTTGTGCGCA	AGATCGCG	CGCGGAGGCTTAC	CGACTTGTGCTGTGCAAGGCTCTC	1012				
QY	901	TGGGCCCCCTGCGCGCTGTCTGTGT	TGCTGTCTTGTGTGGGAGCC	AGCTGAA	CGAGCTC	960				
DB	1013	TGGGCCCCCTGCGCGCTGTCTGTGT	TGCTGTCTTGTGTGGGAGCC	AGCTGAA	CGAGCTC	1072				
QY	961	ACAGCGCAAGGCAATCTCTGTGAGAC	CCCGCTGTGCGAGACA	CCCGAATG	CGCCCTTATGCCCCA	1020				
DB	1073	ACAGCGCAAGGCAATCTCTGTGAGAC	CCCGCTGTGCGAGACA	CCCGAATG	CGCCCTTATGCCCCA	1132				
QY	1021	ACCGGATCCCATCTCTGTGGAGGCTG	CCCAAGTGTGCTCTGATGTGAT	GTGGGCTGTGAGCGAA		1080				
DB	1133	ACCGGATCCCATCTCTGTGGAGGCTG	CCCAAGTGTGCTCTGATGTGAT	GTGGGCTGTGAGCGAA		1192				

QY 1081 GCCAGGGAAGAGGAGGAGACAGAGAGTGTGTTCTGTATGTGTAGACCAACCTTACTTACA 1140
DB 1193 GCCAGGGAAGAGGAGGAGACAGAGAGTGTGTTCTGTATGTGTAGACCAACCTTACTTACA 1252
QY 1141 CGCTCAGCTGCCAAGCAGTGTGATTTAGTTGGGGGTAGCTCCAGCCTTCTCTGCTCTG 1200
DB 1253 CGCTCAGCTGCCAAGCAGTGTGATTTAGTTGGGGGTAGCTCCAGCCTTCTCTGCTCTG 1312
QY 1201 AACTGAGCCAAACCTTCACTGCTTCTTCAAGAGGAGAAAGGAGAGAGCCTGTGTGAGTG 1260
DB 1313 AACTGAGCCAAACCTTCACTGCTTCTTCAAGAGGAGAAAGGAGAGAGCCTGTGTGAGTG 1372
QY 1261 TGTCTGTATACATCTGCTTCTTCTTCAACACATGAGCTTCTGCTGTGCTTATCAG 1320
DB 1373 TGTCTGTATACATCTGCTTCTTCTTCAACACATGAGCTTCTGCTGTGCTTATCAG 1432
QY 1321 GTGCAAGCCCTGCTTCTGCTGTCTGAGAGTACAGCAGTGTGAGCAAGAGAGCAATATTTCC 1380
DB 1433 GTGCAAGCCCTGCTTCTGCTGTCTGAGAGTACAGCAGTGTGAGCAAGAGAGCAATATTTCC 1492
QY 1381 TGTCTCAGAGATGACAACTGCGATCTTGAAGCTGACAACTTTTCAATGACATAGG 1440
DB 1493 TGTCTCAGAGATGACAACTGCGATCTTGAAGCTGACAACTTTTCAATGACATAGG 1552
QY 1441 TCACTGTCTACCTGCTGTGTACCTTTGTACAGATGTGCGCTTCACTGATGTGTGTCTCA 1500
DB 1553 TCACTGTCTACCTGCTGTGTACCTTTGTACAGATGTGCGCTTCACTGATGTGTGTCTCA 1612
QY 1501 GGCACCTGTGTCCAGAGCAATCCCTTTCACAAACAAACAGAGCTGCTTGTATCTTGTGA 1560
DB 1613 GGCACCTGTGTCCAGAGCAATCCCTTTCACAAACAAACAGAGCTGCTTGTATCTTGTGA 1672
QY 1561 CTTTTTCAGAGAAAGGAGGTATCTCTGTGCAAAAGGCTTCAAGCCTTCTCCCTGTCACT 1620
DB 1673 CTTTTTCAGAGAAAGGAGGTATCTCTGTGCAAAAGGCTTCAAGCCTTCTCCCTGTCACT 1732
QY 1621 CAGAGCCCAAGCCAGCTCACTGTGGAAGCTGTGCTCCAGACATCTGTGCTCTTGAAT 1680
DB 1733 CAGAGCCCAAGCCAGCTCACTGTGGAAGCTGTGCTCCAGACATCTGTGCTCTTGAAT 1792
QY 1681 AAGAGATTTCTCTTCCAGGCTTAAGCTGTGGAATTTGGGCAAGATTAAGATCCAACTA 1740
DB 1793 AAGAGATTTCTCTTCCAGGCTTAAGCTGTGGAATTTGGGCAAGATTAAGATCCAACTA 1852
QY 1741 TGAAGCTAGTTCTTGTCTTAATCAAGATGTTCTGTGAATGAAGGCTCCAGCCTGTCAAC 1800
DB 1853 TGAAGCTAGTTCTTGTCTTAATCAAGATGTTCTGTGAATGAAGGCTCCAGCCTGTCAAC 1912
QY 1801 ATGGGGCTTCTGA CTTGAGACCAAGGTTGAGGGAAGATTAAGGAGGCTGTGCTCTGT 1860
DB 1913 ATGGGGCTTCTGA CTTGAGACCAAGGTTGAGGGAAGATTAAGGAGGCTGTGCTCTGT 1972
QY 1861 GGCACCTGTGAAAGTCCAGAGTGGAGCTCTTCTGTGGGACACTTGGGGTCCACAATCCAG 1920
DB 1973 GGCACCTGTGAAAGTCCAGAGTGGAGCTCTTCTGTGGGACACTTGGGGTCCACAATCCAG 2032
QY 1921 GTTCATACCTAGCTTTTGGATACCATGATGATGATTTTACCTGTGCTTAATTAAGGA 1980
DB 2033 GTTCATACCTAGCTTTTGGATACCATGATGATGATTTTACCTGTGCTTAATTAAGGA 2092
QY 1981 GAATTATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2004
DB 2093 GAATTATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2116

RESULT 5
LOCUS AY247738 2488 bp mRNA linear PRI 20-Apr-2003
DEFINITION Homo sapiens TRB3 protein mRNA, complete cds.
ACCESSION AY247738
VERSION AY247738.1 GI:30025661
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2488)
AUTHORS Shan, Y.X. and Yu, L.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2003) School of Life Science, Institute of Genetics, Fudan University, Handan Road 220, Shanghai 200433, PR China

FEATURES
source 1..2488
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 541..1617
 /codon_start=1
 /product="TRB3 protein"
 /protein_id="P404407.1"
 /db_xref="GI:30025662"
 /translation="MRATPLAAPAGLSLRKKRLSDNLDTERPVRQARSQPQLP
PCLLPSPPTAPDRATAVATASRLGPEVYLLPREGGAYOALHCPTEBYCKYPVQ
BALAVLEPYARLPHPKHVAPTEVLATGOLLYAFPTTHGDMSLVSRIRIPERBA
VLFRQATLALHCHQGLVLRDLKLRFPVADREKKLVLENTEDSGVLGPDGLMD
KHACPAVYGEETLSRSASYSGRADVWSLQVALPTMLAGHYPPQDSFVLLFCKIRG
AYALPAGLSAPACLVRLRRERPAERLITATGILHPWLNQDPPLAPTRSHLEAAQ
VVDGJSLDEAREEEDREVLYG"

BASE COUNT 478 a 768 c 746 g 496 t

ORIGIN
Query Match 96.7%; Score 1990.8; DB 9; Length 2488;
Best Local Similarity 99.9%; P-Val: 0;
Matches 1992; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTCTGAGCCCGGCGGCGCCGCGCCAGCGGAAAGAGGCGAGATGCGAGCCACC 60
DB 493 GCTCTGAGCCCGGCGGCGCCGCGCCAGCGGAAAGAGGCGAGATGCGAGCCACC 552
QY 61 CTTCTGAGCTGCTCTGCGGCTTCTCTGTCCAGAGAGAGCGGTGAGTGTGATGACAC 120
DB 553 CTTCTGAGCTGCTCTGCGGCTTCTCTGTCCAGAGAGAGCGGTGAGTGTGATGACAC 612
QY 121 TTGATATACGAGAGGCTCCGTCAGAAACGAGCTCGAAGTGGCCCGCCAGCTGCC 180
DB 613 TTGATATACGAGAGGCTCCGTCAGAAACGAGCTCGAAGTGGCCCGCCAGCTGCC 672
QY 181 CCTGCTGTGTGCGCCCTGAGCCCACTACTGCTCAAGATGTGCAACTGTGTGGCACT 240
DB 673 CCTGCTGTGTGCGCCCTGAGCCCACTACTGCTCAAGATGTGCAACTGTGTGGCACT 732
QY 241 GCTCTCCGCTTGTGAGCCCTTATGTCTCTGTGAGCCCGAGAGGCGGCGGCTTACCG 300
DB 733 GCTCTCCGCTTGTGAGCCCTTATGTCTCTGTGAGCCCGAGAGGCGGCGGCTTACCG 792
QY 301 GCGCTGCACTGCGCTTACAGGCACTGAGTATACCTGTGAGAGGTATACCCGTCAGAGAGCC 360
DB 793 GCGCTGCACTGCGCTTACAGGCACTGAGTATACCTGTGAGAGGTATACCCGTCAGAGAGCC 852
QY 361 CTGGCGTGTGAGGCTTACAGGCGGAGTGGCCCGGCAAGAGATGTGAGTGGCCCACT 420
DB 853 CTGGCGTGTGAGGCTTACAGGCGGAGTGGCCCGGCAAGAGATGTGAGTGGCCCACT 912
QY 421 GAGGTCTGTGCTGTATCCAGAGCTCTCTAGCGCTTTTTCATCTGGAACCATGTGGAGCATG 480
DB 913 GAGGTCTGTGCTGTATCCAGAGCTCTCTAGCGCTTTTTCATCTGGAACCATGTGGAGCATG 972
QY 481 CACAGCTGTGTGGAAGCGGCAACCGTATCCCTGAGCCGAGAGCTGCTGCTTCCCG 540
DB 973 CACAGCTGTGTGGAAGCGGCAACCGTATCCCTGAGCCGAGAGCTGCTGCTTCCCG 1032
QY 541 CAGATGGCCACCGCCCTGTGCGCATGTCAACAGACAGGTCTGTGCTCTGCGTGTATCTCAAG 600
DB 1033 CAGATGGCCACCGCCCTGTGCGCATGTCAACAGACAGGTCTGTGCTCTGCGTGTATCTCAAG 1092

QY 601 CTGTGTGCTTTGTTCTTCTGCTGACCGTGAAGAAAGAAAGCTGTGCTGGAAGAACTGGAG 660
 DB 1093 CTGTGTGCTTTGTTCTTCTGCTGACCGTGAAGAAAGAAAGCTGTGCTGGAAGAACTGGAG 1152
 QY 661 GACTCTGCGGCTGACCTGAGGAGCAAGATGATTCCTGCTGGGACAAGCAAGCGCTGCGAGCC 720
 DB 1153 GACTCTGCGGCTGACCTGAGGAGCAAGATGATTCCTGCTGGGACAAGCAAGCGCTGCGAGCC 1212
 QY 721 TACGTGGAGCTTGAAGATTAAGTCAAGCTCAAGGAGCTTCACTGCGGAGAAAGCAAGCGAGTTC 780
 DB 1213 TACGTGGAGCTTGAAGATTAAGTCAAGCTCAAGGAGCTTCACTGCGGAGAAAGCAAGCGAGTTC 1272
 QY 781 TGAAGCTGAGGAGCTGAGGAGCTTCAAGCTGCTGAGCGGAGCACTAACCTTTTCAAGAGACTGC 840
 DB 1273 TGAAGCTGAGGAGCTGAGGAGCTTCAAGCTGCTGAGCGGAGCACTAACCTTTTCAAGAGACTGC 1332
 QY 841 GAGGCTGCTGCTCTTCTGAGCAAGATTCGAGCGGAGGAGCTCAAGGCTTTCGAGGAGCTTC 900
 DB 1333 GAGGCTGCTGCTCTTCTGAGCAAGATTCGAGCGGAGGAGCTCAAGGCTTTCGAGGAGCTTC 1392
 QY 901 TCGGAGCTGCTGCTCTTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 1393 TCGGAGCTGCTGCTCTTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1452
 QY 961 ACAGCCACAGGAGCTTCTCTGCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB 1453 ACAGCCACAGGAGCTTCTCTGCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
 QY 1021 ACCGAGTCCCATCTCTGAGGAGGCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 DB 1513 ACCGAGTCCCATCTCTGAGGAGGCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
 QY 1081 GCCAGGAG 1140
 DB 1573 GCCAGGAG 1632
 QY 1141 CGCTCAGCTGCGCAAGTGAATGATTTGGGGAGTACCTCAAGCTTCTCTGCTGCTGCTGCTGCTGCT 1200
 DB 1633 CGCTCAGCTGCGCAAGTGAATGATTTGGGGAGTACCTCAAGCTTCTCTGCTGCTGCTGCTGCTGCT 1692
 QY 1201 AACTGAGCCAAACCTTCAAGTCTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 DB 1693 AACTGAGCCAAACCTTCAAGTCTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1752
 QY 1261 TGCTGTGTACACATCTGCTTTGTTTCCACACACATGCAAGTCTCTGCTGCTGCTGCTGCTGCT 1320
 DB 1753 TGCTGTGTACACATCTGCTTTGTTTCCACACACATGCAAGTCTCTGCTGCTGCTGCTGCTGCT 1812
 QY 1321 GTGCAAGCCCTGTTCTCGGAGTGAAGAGTACAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 DB 1813 GTGCAAGCCCTGTTCTCGGAGTGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1872
 QY 1381 TGCTCAGAGAGTGAACAACTGGAGCTTGAAGTGAACAACTTTTCCATGACATAGG 1440
 DB 1873 TGCTCAGAGAGTGAACAACTGGAGCTTGAAGTGAACAACTTTTCCATGACATAGG 1932
 QY 1441 TCACTGTCTACATCTGAGGAGTCACTTTGTAACAGTGTGCGGCTTCACTGATGCTGCTGCTCA 1500
 DB 1933 TCACTGTCTACATCTGAGGAGTCACTTTGTAACAGTGTGCGGCTTCACTGATGCTGCTGCTCA 1992
 QY 1501 GGCACCTCTGTCAAGAGCAATCCCTTTCAAAACAAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 DB 1993 GGCACCTCTGTCAAGAGCAATCCCTTTCAAAACAAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2052
 QY 1561 CTTTTCAGAGAGAGAGAGTATCCCTGTGCAAAAGCTTCAGAGCTTCTCCCTGCAACT 1620
 DB 2053 CTTTTCAGAGAGAGAGAGTATCCCTGTGCAAAAGCTTCAGAGCTTCTCCCTGCAACT 2112
 QY 1621 CAGGAGCCAGGAGCTCACTCTGAGAGAGTGTCCAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
 DB 2113 CAGGAGCCAGGAGCTCACTCTGAGAGAGTGTCCAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2172

QY 1681 AAGAGATTCTCTTCCAGGCTTAAGCTGAGATTGCGGCGAGAGATTAAGATCCAAACTA 1740
 DB 2173 AAGAGATTCTCTTCCAGGCTTAAGCTGAGATTGCGGCGAGAGATTAAGATCCAAACTA 2232
 QY 1741 TGAAGCTAGTTCTTGTCTTAAGTCAAGCTGTTCTGGAATGAGAGGCTGAGGCTGCAACC 1800
 DB 2233 TGAAGCTAGTTCTTGTCTTAAGTCAAGCTGTTCTGGAATGAGAGGCTGAGGCTGCAACC 2292
 QY 1801 ATGGAGGCTTCTGACCTGAGCAACAGGTTGAGGAGCAAGATTAGCGAGGCTGCTGCTGCT 1860
 DB 2293 ATGGAGGCTTCTGACCTGAGCAACAGGTTGAGGAGCAAGATTAGCGAGGCTGCTGCTGCTGCT 2352
 QY 1861 GGCACCTGGAAGATCCAGGTGAGACTTCTGAGGAGCACTTGGGGGTGCAAAATCCAG 1920
 DB 2353 GGCACCTGGAAGATCCAGGTGAGACTTCTGAGGAGCACTTGGGGGTGCAAAATCCAG 2412
 QY 1921 GTCCATACCTAGATTGATACATGATGATGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
 DB 2413 GTCCATACCTAGATTGATACATGATGATGATGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2472
 QY 1981 GAATTATGAATAA 1994
 DB 2473 GAATTATGAATAA 2486

RESULT 6
 HS110367
 LOCUS
 DEFINITION
 Human DNA sequence from clone RP5-110367 on chromosome 20p12.2-13. Contains up to three novel genes, the gene for a novel protein similar to mouse VMP, the gene for a novel protein kinase domain containing protein similar to phosphoprotein C8FW and rat NPK, and the SOX22 gene for SRY (sex-determining region Y)-box 22. Contains five CpG islands, ESTs, STSs and GSSs, complete sequence.

ACCESSION
 AL034548
 VERSION
 AL034548.25 GI:7263904
 KEYWORDS
 HMG; CPG Island; NPK; protein kinase; SOX22; SRY; VMP.

SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens

REFERENCE
 AUTHORS
 Bakkayola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Primates; Catarrhini; Homidae; Homo.

JOURNAL
 Direct Submission
 Submitted (05-FEB-2001) Sanger Centre, Hinxton, Cambridgehire, CB10 1SA, UK. E-mail enquiries: humque@anger.ac.uk
 requests: clonerequest@anger.ac.uk
 On Mar 19, 2000 this sequence version replaced gi:5541861.

COMMENT
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TrEMBL; Wp; WORMBEP; Information on the WORMBEP database can be found at http://www.sanger.ac.uk/Project/C_elegans/wormbep This sequence is the entire insert of clone RP5-110367 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/RGP/Chr20> RP5-110367 is from the library RPCI-5 constructed by the group of Pleters de Jong. For further details see <http://www.chori.org/bacpac/home.htm>


```
VECTOR: pcvpac2.
FEATURES
  source
    Location/Qualifiers
      1..153170
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="20"
        /map="p12.2-13"
        /clone="RP5-110JG7"
        /clone_11b="RPC1-5"
      1..61
        /note="Alu repeat: matches 2..62 of consensus"
      79..129
        /note="Alu repeat: matches 6221..6271 of consensus"
      130..173
        /note="L1M9 repeat: matches 2 mer ag 79% conserved"
      175..873
        /note="L1M9 repeat: matches 5519..6217 of consensus"
      902..1109
        /note="L1M5 repeat: matches 7715..7921 of consensus"
      1110..1480
        /note="L1M4 repeat: matches 1..371 of consensus"
      1502..1733
        /note="match: STR: Em:266755"
      1734..1888
        /note="L1M1H repeat: matches 387..547 of consensus"
      1919..2101
        /note="L1M1H repeat: matches 115..298 of consensus"
      complement(join(2146..2243,126791..126834))
        /note="match: GSS: Em:B59651"
      2154..2193
        /note="20 copies 2 mer aa 85% conserved"
      complement(join(2194..2243,126792..126871))
        /note="match: GSS: Em:AQ377604"
      2639..2917
        /note="L1M4 repeat: matches 2875..3081 of consensus"
      2938..3062
        /note="L1M4 repeat: matches 3120..3249 of consensus"
      3136..3430
        /note="L1M5 repeat: matches 1..295 of consensus"
      3482..3777
        /note="L1M5 repeat: matches 1..296 of consensus"
      3811..4126
        /note="L1M5 repeat: matches 1..312 of consensus"
      4310..4608
        /note="L1M2 repeat: matches 5861..6167 of consensus"
      4609..5026
        /note="L1M2 repeat: matches 1..424 of consensus"
      5027..5200
        /note="L1M2 repeat: matches 5684..5861 of consensus"
      5201..5501
        /note="L1M2 repeat: matches 1..307 of consensus"
      5502..5727
        /note="L1M2 repeat: matches 5451..5684 of consensus"
      5798..5865
        /note="34 copies 2 mer aa 66% conserved"
      6056..6643
        /note="MER67C repeat: matches 122..710 of consensus"
      6648..6707
        /note="30 copies 2 mer ca 88% conserved"
      6710..6824
        /note="MER50 repeat: matches 10..134 of consensus"
      6825..7124
        /note="L1M5 repeat: matches 7..306 of consensus"
      7126..7415
        /note="L1M4 repeat: matches 5619..5919 of consensus"
      7419..7569
        /note="L1M5 repeat: matches 148..298 of consensus"
      7576..7889
        /note="match: GSS: Em:AQ180303"
      complement(7590..7894)
        /note="match: GSS: Em:AQ01808"
      7862..8140
        /note="match: GSS: Em:AQ123077"
      8019..8317
        /note="match: GSS: Em:AQ746749"
      8019..8151
        /note="match: GSS: Em:AQ617736"
      8152..8316
        /note="MER4B repeat: matches 416..574 of consensus"
      8349..8643
        /note="MER4D repeat: matches 362..689 of consensus"
      8644..8942
        /note="L1M5 repeat: matches 1..299 of consensus"
      8943..9113
        /note="MER4D repeat: matches 194..362 of consensus"
      9113..9201
        /note="L1M2 repeat: matches 377..450 of consensus"
      9226..9311
        /note="L1M4 repeat: matches 5775..5862 of consensus"
      9312..9651
        /note="MER1B repeat: matches 1..337 of consensus"
      9653..9750
        /note="L1M5 repeat: matches 5856..5947 of consensus"
      9751..10047
        /note="L1M5 repeat: matches 1..298 of consensus"
      10048..10369
        /note="L1M5 repeat: matches 5947..6300 of consensus"
      10488..10529
        /note="21 copies 2 mer at 88% conserved"
      10532..10653
        /note="L1M4 repeat: matches 7673..7800 of consensus"
      10687..10752
        /note="13 copies 2 mer at 68% conserved"
      11681..11742
        /note="31 copies 2 mer ca 71% conserved"
      13519..13972
        /note="match: GSS: Em:AQ224309"
      14208..14554
        /note="L1M1B repeat: matches 110..462 of consensus"
      15554..15872
        /note="L1M1 repeat: matches 5615..5910 of consensus"
      15873..16174
        /note="L1M5 repeat: matches 16..313 of consensus"
      16175..16382
        /note="L1M1 repeat: matches 5400..5615 of consensus"
      16383..16423
        /note="L1M2 repeat: matches 5893..5933 of consensus"
      16595..16894
        /note="L1M2 repeat: matches 1..302 of consensus"
      16895..17021
        /note="L1M1 repeat: matches 6196..6318 of consensus"
      17155..17587
        /note="L1M2 repeat: matches 5289..5733 of consensus"
      17587..18157
        /note="L1M1 repeat: matches -1389..-790 of consensus"
      18158..18195
        /note="19 copies 2 mer tg 97% conserved"
      18197..18230
        /note="L1M5 repeat: matches 165..197 of consensus"
      18231..18823
        /note="MER52A repeat: matches 1..672 of consensus"
      complement(18827..19180)
        /note="match: GSS: Em:AQ59398"
      19689..19998
        /note="L1M5 repeat: matches 1..309 of consensus"
      20287..20543
        /note="L1M5 repeat: matches 2..307 of consensus"
      20552..20609
        /note="L1M1 repeat: matches 255..318 of consensus"
      20592..20756
        /note="L1M1 repeat: matches 256..421 of consensus"
      20895..21012
        /note="L1M5 repeat: matches -1128..-1008 of consensus"
      21009..21168
        /note="L1M5 repeat: matches -1537..-1377 of consensus"
```



```

repeat_region 21192..21277
/note="LIM47 repeat: matches 5864..5949 of consensus"
repeat_region 21280..21627
/note="L2 repeat: matches 2250..2617 of consensus"
misc_feature 21628..21884
/note="match: G8S: Em:B41180"
misc_feature 21628..21744
/note="match: G8S: Em:A0058160"

Query Match 66.4%; Score 1367; DB 9; Length 153170;
Best Local Similarity 98.9%; Pred. No. 1.8e-222;
Matches 1376; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 604 TGTGCTTTTGTCTTCTGCTGACCGTGAAGAAAGCTGGTCTGAGAACCTGAGAGC 663
DB 150847 TGACCCCTTCTGTTTCTCCCCATGTCACAGAGAAAGCTGGTCTGAGAACCTGAGAGC 150906
QY 664 TCCGCGGTGCTGACCTGGAGCCAGATGATTTCCCTGAGGACAAAGCAGGAGCCAGCTTAC 723
DB 150907 TCCGCGGTGCTGACCTGGAGCCAGATGATTTCCCTGAGGACAAAGCAGGAGCCAGCTTAC 150966
QY 724 GTGGAGCTGAGATATCACTCACTCAAGGAGCTCATATCTCGGCAAGAGCCAGATGTCTGG 783
DB 150967 GTGGAGCTGAGATATCACTCACTCAAGGAGCTCATATCTCGGCAAGAGCCAGATGTCTGG 151026
QY 784 AGCCTGGAGCTGGCTGCTTCTTCAACATGCTGGCCGCACTAACCCCTTCCAGAACTCGAG 843
DB 151027 AGCCTGGAGCTGGCTGCTTCTTCAACATGCTGGCCGCACTAACCCCTTCCAGAACTCGAG 151086
QY 844 CTTGCTGCTGCTTCCGCAAGATCGGCGGAGGAGCTAACGCTTGGCCCTGAGGCTCTCG 903
DB 151087 CTTGCTGCTGCTTCCGCAAGATCGGCGGAGGAGCTAACGCTTGGCCCTGAGGCTCTCG 151146
QY 904 GCCCTGCGCCGCTGTCTGTTCTGCTGCTCTCTTGTGCGAGAGCCAGCTGAACGAGCTCA 963
DB 151147 GCCCTGCGCCGCTGTCTGTTCTGCTGCTCTCTTGTGCGAGAGCCAGCTGAACGAGCTCA 151206
QY 964 GCCAGAGGATCTCTCTGACACCCCTGCTGAGACAGAGCCGAGTCCCTTAAGCCCAACC 1023
DB 151207 GCCAGAGGATCTCTCTGACACCCCTGCTGAGACAGAGCCGAGTCCCTTAAGCCCAACC 151266
QY 1024 CGATCCCATCTCTGAGAGGCTGCCAGAGTGGTCCCTGATGGAAGCTGGAGCTGAGAGAGC 1083
DB 151267 CGATCCCATCTCTGAGAGGCTGCCAGAGTGGTCCCTGATGGAAGCTGGAGCTGAGAGAGC 151326
QY 1084 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1143
DB 151327 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151386
QY 1144 TCAAGCTGCGCAAGTGAATTTGAGTTTGGGGTAACTCAAGCCCTTCTCGGCTCTGAGC 1203
DB 151387 TCAAGCTGCGCAAGTGAATTTGAGTTTGGGGTAACTCAAGCCCTTCTCGGCTCTGAGC 151446
QY 1204 TGAGCCCAAACTTCAAGTCCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1263
DB 151447 TGAGCCCAAACTTCAAGTCCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151506
QY 1264 TGTGTACACATCTGCTTTGTTTCCACACATGACATGCTTGTGGTGGCTTATCAGAGT 1323
DB 151507 TGTGTACACATCTGCTTTGTTTCCACACATGACATGCTTGTGGTGGCTTATCAGAGT 151566
QY 1324 CCAAGCCCTGTTCTCGGCTGCTGGAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1383
DB 151567 CCAAGCCCTGTTCTCGGCTGCTGGAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151626
QY 1384 TCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1443
DB 151627 TCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151686
QY 1444 CTGTCTACATCTGGGTACATTTTGTACAGAGTGGGCTTCACTGATAGTGTGTCTCAGGC 1503
DB 151687 CTGTCTACATCTGGGTACATTTTGTACAGAGTGGGCTTCACTGATAGTGTGTCTCAGGC 151746

```

```

QY 1504 ACCTGTGTCCAGAGACAAATCCCTTTTTCACAAACAAACAGCTGCTTTGTATCTTGTACT 1563
DB 151747 ACCTGTGTCCAGAGACAAATCCCTTTTTCACAAACAAACAGCTGCTTTGTATCTTGTACT 151806
QY 1564 TTTTCAGAGAAAGGAGATATCCCTGTGACCAAGAGCTCCAGAGCTCTCCCTGCAACTCAG 1623
DB 151807 TTTTCAGAGAAAGGAGATATCCCTGTGACCAAGAGCTCCAGAGCTCTCCCTGCAACTCAG 151866
QY 1624 GACCCAAAGCCAGCTCACTGTGGAACCTGTGTTCCAGCATCTGTCTCTTGTATTAG 1683
DB 151867 GACCCAAAGCCAGCTCACTGTGGAACCTGTGTTCCAGCATCTGTCTCTTGTATTAG 151926
QY 1684 AGATTCTCTTCCAGAGCTTAAAGCTGGAGATTTTGGCCAGAGATTAAGATCAAACTATGA 1743
DB 151927 AGATTCTCTTCCAGAGCTTAAAGCTGGAGATTTTGGCCAGAGATTAAGATCAAACTATGA 151986
QY 1744 GGCATGTTCTTGTCTTAACTCAAGATGTTCTGGAATGAGAGGTCCAGAGCTGTCAACATG 1803
DB 151987 GGCATGTTCTTGTCTTAACTCAAGATGTTCTGGAATGAGAGGTCCAGAGCTGTCAACATG 152046
QY 1804 GGGCTTCTGACCTGAGACCAAGGTTGAGGAGACAGATTAAGCAGAGTCTGTCTGTGC 1863
DB 152047 GGGCTTCTGACCTGAGACCAAGGTTGAGGAGACAGATTAAGCAGAGTCTGTCTGTGC 152106
QY 1864 CACCTGGAAGATCCCAAGTGGAGACTCTTCTGGGAGACATTTGGGCTCCAAATCCCAAGTC 1923
DB 152107 CACCTGGAAGATCCCAAGTGGAGACTCTTCTGGGAGACATTTGGGCTCCAAATCCCAAGTC 152166
QY 1924 CATACTCTAGTTTGTGATACATAGATATGATATGTTACTGTGCTTAATTAAGAGAA 1983
DB 152167 CATACTCTAGTTTGTGATACATAGATATGATATGTTACTGTGCTTAATTAAGAGAA 152226
QY 1984 TTATGAATTA 1994
DB 152227 TTATGAATTA 152237

RESULT 7
AX166518 1077 bp DNA linear PAT 22-JUN-2001
LOCUS AX166518
DEFINITION Sequence 9 from Patent WO0138503.
ACCESSION AX166518
VERSION AX166518.1 GI:14546863
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1
Plozman, G.D., Whyte, D., Manning, G.S., Sudarshanam, S.S., Martinez, R.,
Pianagan, P. and Clary, D.S.
Novel human protein kinases and protein kinase-like enzymes
Patent: WO 0138503-A 9 31-MAY-2001,
Sugen, Inc. (US)
FEATURES
Source 1..1077
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 174 a 373 c 325 g 205 t
ORIGIN
Query Match 52.2%; Score 1073.8; DB 6; Length 1077;
Best Local Similarity 99.8%; Pred. No. 1.9e-196;
Matches 1075; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 49 ATGCGAGCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 108
DB 1 ATGCGAGCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
QY 109 TTGATGACAACTTATGATACCGAGCTCCCGTCCAGAAAAGAGCTGAGAGGAGGAGGAG 168
DB 61 TTGATGACAACTTATGATACCGAGCTCCCGTCCAGAAAAGAGCTGAGAGGAGGAGGAGGAG 120

```

QY	169	CCGAGACTGACCCCTGCTGTGGCCCTGAGGCCACTATGCTCTCCAGATGTGCACT	228
Db	121	CCGAGACTGACCCCTGCTGTGGCCCTGAGGCCACTATGCTCTCCAGATGTGCACT	180
QY	229	GCTGTGGCACTGCTCCCGTCTTGGGCGTAATGTCCTCTTGAGACCCGAGAGGGCGGG	288
Db	181	GCTGTGGCACTGCTCCCGTCTTGGGCGTAATGTCCTCTTGAGACCCGAGAGGGCGGG	240
QY	289	CGGGCTTACGGGGCCCTGCACTGGCCCTTACAGGCACTGATATACCTGGAAGTGTACCCC	348
Db	241	CGGGCTTACAGGGCCCTGCACTGGCCCTTACAGGCACTGATATACCTGGAAGTGTACCCC	300
QY	349	GTCCAGGAGGCCCTGAGCGCTGTGGAGGCCCTACGGCGGCGTGGCCCGCACAGCATGTG	408
Db	301	GTCCAGGAGGCCCTGAGCGCTGTGGAGGCCCTTATGCGCGGCTGGCCCGCACAGCATGTG	360
QY	409	GCTCGGCCCACTGAGGTCTGTGGCTGTGTACCGAGCTCTTACGCGCTTTTTCATCTGAGCC	468
Db	361	GCTCGGCCCACTGAGGTCTGTGGCTGTGTACCGAGCTCTTACGCGCTTTTTCATCTGAGCC	420
QY	469	CATGGGAGCATGCAAGCTGTGTGTGGAAAGCGGCACCGTATTCCTGAGCCTGAGGCTGCC	528
Db	421	CATGGGAGCATGCAAGCTGTGTGTGGAAAGCGGCACCGTATTCCTGAGCCTGAGGCTGCC	480
QY	529	GTGCTCTTCCGCGCAGATGGCCACGCGCCCTGGGGCATCTGTACACAGCAGAGTCTGGTCTG	588
Db	481	GTGCTCTTCCGCGCAGATGGCCACGCGCCCTGGGGCATCTGTACACAGCAGAGTCTGGTCTG	540
QY	589	CGTATCTCAAGCTGTGTGTGGCTTTGTCTTTCGCTGTACCGTGAAGAGAAAGCTGTGTCTG	648
Db	541	CGTATCTCAAGCTGTGTGTGGCTTTGTCTTTCGCTGTACCGTGAAGAGAAAGCTGTGTCTG	600
QY	649	GAGAACTGTGAGAGACTCTGTGTGTGTACTGTGGCCAGATGATTCCTGTGGACAGAGAC	708
Db	601	GAGAACTGTGAGAGACTCTGTGTGTGTACTGTGGCCAGATGATTCCTGTGGACAGAGAC	660
QY	709	GCGTGGCCAGGCTACGTGGGAGCTGAGATATCTCAGCTCAAGGGCCCTCATATCTGGGCAAG	768
Db	661	GCGTGGCCAGGCTACGTGGGAGCTGAGATATCTCAGCTCAAGGGCCCTCATATCTGGGCAAG	720
QY	769	GCAGCCGATGTCTGTAGAGCTGTGGGCGTGGCGCTTTCACATATGTGGGCGCATCTACCCC	828
Db	721	GCAGCCGATGTCTGTAGAGCTGTGGGCGTGGCGCTTTCACATATGTGGGCGCATCTACCCC	780
QY	829	TTCCAGAGACTGTGAGGCTGTCTGTCTTTCGCGCAAGATTCGCGCGGGGCTTACGCTTGG	888
Db	781	TTCCAGAGACTGTGAGGCTGTCTGTCTTTCGCGCAAGATTCGCGCGGGGCTTACGCTTGG	840
QY	889	CTTGCAGGCGCTCTGGGCGCCCGTGGCGCGTGTGTTGTGCTGCTCCCTTCGTCTGGAGGCA	948
Db	841	CTTGCAGGCGCTCTGGGCGCCCGTGGCGCGTGTGTTGTGCTGCTCCCTTCGTCTGGAGGCA	900
QY	949	GCTTAAAGGCTCACAGGCGCACAGGCACTCTCCCTGTGCACCCCTGTGGCTGTGACAGGACCCGATG	1008
Db	901	GCTTAAAGGCTCACAGGCGCACAGGCACTCTCTGTGCACCCCTGTGGCTGTGACAGGACCCGATG	960
QY	1009	CCCTTAAAGCCCAACCCGATTCCTATCTCTGGAGGCTGTGCCAGGTGTCTCTGATGAGCTG	1068
Db	961	CCCTTAAAGCCCAACCCGATTCCTATCTCTGGAGGCTGTGCCAGGTGTCTCTGATGAGCTG	1020
QY	1069	GGGCTTGACCAAGCAGAGGAGAGAGAGGAGACAGAGAAAGTGTCTGTATGGCTAG	1125
Db	1021	GGGCTTGACCAAGCAGAGGAGAGAGAGGAGACAGAGAAAGTGTCTGTATGGCTAG	1077

RESULT 8	AX572896	1083 bp	DNA	1linear	PAT 29-NOV-2002
LOCUS	AX572896				
DEFINITION	Sequence 3 from Patent WO0205743.				
ACCESSION	AX572896				
VERSION	AX572896.1	GI:26004968			
KEYWORDS	.				

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1
AUTHORS	Dower, S. and Quane, R. E.
TITLE	Mammalian tribbles signaling pathways and methods and reagents related thereto
JOURNAL	Patent: WO 0203743-A 3 11-JUL-2002;
FEATURES	Interleukin Genetics, Inc. (US)
SOURCE	Location/Qualifiers
	1..1083
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
BASE COUNT	176 a 374 c 327 g 206 t
ORIGIN	
Query Match	51.2%; Score 1054.6; DB 6; Length 1083;
Best Local Similarity	99.1%; Pred. No. 9.6e-193;
Matches 1073; Conservative	0; Mismatches 4; Indels 6; Gaps 1;
QY	49 ATGGAGACCAACCCCTCTGGCTGCTCCCGGGGTTCCCTGACGAAGAAGCCGTTGGAG 108
DB	1 ATGGAGACCAACCCCTCTGGCTGCTCCCGGGGTTCCCTGACGAAGAAGCCGTTGGAG 60
QY	109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCAGAAAGAGCTCGAAGTGGCCCCAG 168
DB	61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCAGAAAGAGCTCGAAGTGGCCCCAG 120
QY	169 CCCAGACTGCCCCCTGCTGCTTTGCTCCCTGAGCCCACTACTGCTCCAGATCGTGCACT 228
DB	121 CCCAGACTGCCCCCTGCTGCTTTGCTCCCTGAGCCCACTACTGCTCCAGATCGTGCACT 180
QY	229 GCTGTGGCCACTGCTCCCGCGCTTTGGGCGCTTAAGTCTCTCTGAGCGCCGAGAGGGGGGG 288
DB	181 GCTGTGGCCACTGCTCCCGCTTTGGGCGCTTAAGTCTCTCTGAGCGCCGAGAGGGGGGG 240
QY	289 CGGGCTACACCGGGCCCTGCACTGCCCTCAACAGGCACTGATATACCTGCAAGTGTACCCC 348
DB	241 CGGGCTACACAGGCGCCCTGCACTGCCCTCAACAGGCACTGATATACCTGCAAGTGTACCCC 300
QY	349 GTCCAGAGAACCTGCGCGTGTGAGAGCCCTAAGCGCGCTGCCCGCACAGCATGTG 408
DB	301 GTCCAGAGAACCTGCGCGTGTGAGAGCCCTAAGCGCGGATGCGCCCGCACAGCATGTG 360
QY	409 GCTCGAGCCCACTGAGGCTCTGGCGTGAACCAAGCTCTCTACAGGCTTTTTCACCTGGAGC 468
DB	361 GCTCGAGCCCACTGAGGCTCTGGCGTGAACCAAGCTCTCTACAGGCTTTTTCACCTGGAGC 420
QY	469 CATTGGAGCACTGACAGCCTGTGTGCGAAGCCGCAACCGTATCCCTGAGCCTGAGGCTGCC 528
DB	421 CATTGGAGCACTGACAGCCTGTGTGCGAAGCCGCAACCGTATCCCTGAGCCTGAGGCTGCC 480
QY	529 GTGCTCTTCCGCCCAAGATGAGGCAACCGCCCTGAGCGCACTGTCAACAGCAACGATGTGTCTG 588
DB	481 GTGCTCTTCCGCCCAAGATGAGGCAACCGCCCTGAGCGCACTGTCAACAGCAACGATGTGTCTG 540
QY	589 CGTATCTCAAGCTGTGTGCTTTGTCTTGTGC-----TGACCGTGTGAGAGAAAGACTG 642
DB	541 CGTATCTCAAGCTGTGTGTGCTTTGTCTTGTGCTGACCGTATCCCTGAGAGAAAGACTG 600
QY	643 GTGCTGAGAACTTGAGAGACTCTGTGTGTCTGACTGATGAGCCAGATGATTCCTGTGTGGAC 702
DB	601 GTGCTGAGAACTTGAGAGACTCTGTGTGTCTGACTGATGAGCCAGATGATTCCTGTGTGGAC 660
QY	703 AAGCAACGAGCCAGCTTACGTTAGGAGACCTGAGATTACTGACGTCAAGGGCTTAATATCG 762
DB	661 AAGCAACGAGCCAGCTTACGTTAGGAGACCTGAGATTACTGACGTCAAGGGCTTAATATCG 720
QY	763 GGCAGAGCAGCCAGATCTTGAAGCCTGTGGCGTGGCGCTTTTCACTCAATGTCTGGCCGAC 822
DB	721 GGCAGAGCAGCCAGATCTTGAAGCCTGTGGCGTGGCGCTTTTCACTCAATGTCTGGCCGAC 780

QY 823 TACCCCTTCCAGAGACTCGAGCCTGTCTCTCTTCGGAAGATCCGCGGGGCTTAC 882
DB 781 TACCCCTTCCAGAGACTCGAGCCTGTCTCTCTTCGGAAGATCCGCGGGGCTTAC 840
QY 883 GCTTGTGCTGAGAGCCTCTGCGCCCTGCGCTGTCTGTGTTGCTGCTGCTCTTCTGCGG 942
DB 841 GCTTGTGCTGAGAGCCTCTGCGCCCTGCGCTGTCTGTGTTGCTGCTGCTCTTCTGCGG 900
QY 943 GAGCCAGCTGAACGGCTCAGAGCCAGAGCATCTCTGCAACCCCTGCTGCAAGAG 1002
DB 901 GAGCCAGCTGAACGGCTCAGAGCCAGAGCATCTCTGCAACCCCTGCTGCAAGAG 960
QY 1003 CCGATGCTTAAAGCCCAACCCGATCCCATCTCTGAGAGCTGCGCAAGTGTCTGAT 1062
DB 961 CCGATGCTTAAAGCCCAACCCGATCCCATCTCTGAGAGCTGCGCAAGTGTCTGAT 1020
QY 1063 GGACTGGGGCTGAGCAGAGCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122
DB 1021 GGACTGGGGCTGAGCAGAGCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1123 TAG 1125
DB 1081 TAG 1083

RESULT 9
AF250311 1083 bp mRNA linear PRI 01-JUN-2001
LOCUS Homo sapiens SKIP3 mRNA, complete cds.
DEFINITION AF250311 AF250311.1 GI:14276268
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
1 (bases 1 to 1083)
Kiss-Toth, E., Wylie, D.H., Ovarstrom, E.E. and Dower, S.K.
Identification of pro-inflammatory cytokine signalling network
components by transcription expression screening
Unpublished
2 (bases 1 to 1083)
Kiss-Toth, E., Wylie, D.H., Ovarstrom, E.E. and Dower, S.K.
Direct Submission
Submitted (29-MAR-2000) Division of Molecular & Genetic Medicine,
University of Sheffield, Royal Hallamshire Hospital, Floor M,
Sheffield S10-2JF, UK
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1083
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1083
/codon_start=1
/product="SKIP3"
/protein_id="AAK58175.1"
/db_xref="GI:14276269"
/translation="MRATPLAAGSLSRKKRLDNLDTERRPVOKASGPOPLP
PCLLPSPAPRATVATASRLGPTVLLPEEGRAYALCPSTETCYKVPVQ
BALALBPVAPVPHKVARPTVLGTLVLAFTTHDMSLVSRIRIEBPA
VLPROMATLALHCHGIVRLDKCRFVADRDREKKLVLENDSCVTLPDPSL
MDKACPAYVPEILSRASYSKADWVSLGVALPTMLAGHPDPSSEVYLFGKIR
KGAVALPAGLSAPARCLVRLRRRPRRLTAGILHLPWLRDPPMLAFTSLMWA
AQPVPDLGIDREBEREGREVLYG"

BAGE COUNT 176 a 374 c 327 g 206 t
ORIGIN

Query Match 51.2%; Score 1054.6; DB 9; Length 1083;
Best Local Similarity 99.1%; Pred. No. 9.6e-193;
Matches 1073; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 49 ATGCGAGCACCCCTGTGCTGCTCTGCGGGTTCCCTGTCCAGAAAGACGGATTGAG 108

DB 1 ATGCGAGCACCCCTGTGCTGCTCTGCGGGTTCCCTGTCCAGAAAGACGGATTGAG 60
QY 109 TTGATATGACAACTTAGATACCGAGCGTCCGCTCAGAAACAGACTCGAGTGGCCCCAG 168
DB 61 TTGATATGACAACTTAGATACCGAGCGTCCGCTCAGAAACAGACTCGAGTGGCCCCAG 120
QY 169 CCCAAGCTGCCCCCTGCTGTGTTGCCCTGAGGCCACTTACTGTCTCAAGATTCGTGA 228
DB 121 CCCAAGCTGCCCCCTGCTGTGTTGCCCTGAGGCCACTTACTGTCTCAAGATTCGTGA 180
QY 229 GCTGTGGCCATGCGCTCCCGCTTTGGGGCCCTAGTGTCTCTGTGAGCCCGAGAGGGGG 288
DB 181 GCTGTGGCCATGCGCTCCCGCTTTGGGGCCCTAGTGTCTCTGTGAGCCCGAGAGGGGG 240
QY 289 CCGAGCTTACCGGGCCCTGCACTGCCCTTACAGAGCACTAGATTACTTCAAGTGTACCC 348
DB 241 CCGAGCTTACCGGGCCCTGCACTGCCCTTACAGAGCACTAGATTACTTCAAGTGTACCC 300
QY 349 GTCAGAGAAAGCCCTGAGCCCTGCTGAGAGCCCTTACAGAGCCCTGCTGAGAGCC 408
DB 301 GTCAGAGAAAGCCCTGAGCCCTGCTGAGAGCCCTTACAGAGCCCTGCTGAGAGCC 360
QY 409 GCTCGGCCCCATGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 468
DB 361 GCTCGGCCCCATGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 469 CATGGGAGCATGACAGAGCTGTGTCGAAAGCCGACACCTGATCCCTGAGCTGAGCTGCC 528
DB 421 CATGGGAGCATGACAGAGCTGTGTCGAAAGCCGACACCTGATCCCTGAGCTGAGCTGCC 480
QY 529 GTGCTCTTCCGAGATGAGCAGCCGCTGAGCCGCTGACAGCAAGATGCTGCTGCTG 588
DB 481 GTGCTCTTCCGAGATGAGCAGCCGCTGAGCCGCTGACAGCAAGATGCTGCTGCTG 540
QY 589 CGTGAATCTCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
DB 541 CGTGAATCTCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 643 GTGCTGAGAGAACTGAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
DB 601 GTGCTGAGAGAACTGAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 703 AAGCAGCGTGTCCGAGCTTACGTGAGCACTGAGATCTCAAGCTTACCGAGCTTACTG 762
DB 661 AAGCAGCGTGTCCGAGCTTACGTGAGCACTGAGATCTCAAGCTTACCGAGCTTACTG 720
QY 763 GGCAGAGCGAGATGCTGAGAGCTGAGGAGGAGGCTTTCACAGCTGAGCTGAGCTGAG 822
DB 721 GGCAGAGCGAGATGCTGAGAGCTGAGGAGGAGGCTTTCACAGCTGAGCTGAGCTGAG 780
QY 823 TACCCCTTCCAGAGACTCGAGCCTGTCTCTCTTCGGAAGATCCGCGGGGCTTAC 882
DB 781 TACCCCTTCCAGAGACTCGAGCCTGTCTCTCTTCGGAAGATCCGCGGGGCTTAC 840
QY 883 GCTTGTGCTGAGAGCCTCTGCGCCCTGCGCTGTCTGTGTTGCTGCTGCTCTTCTGCGG 942
DB 841 GCTTGTGCTGAGAGCCTCTGCGCCCTGCGCTGTCTGTGTTGCTGCTGCTCTTCTGCGG 900
QY 943 GAGCCAGCTGAACGGCTCAGAGCCAGAGCATCTCTGCAACCCCTGCTGCAAGAG 1002
DB 901 GAGCCAGCTGAACGGCTCAGAGCCAGAGCATCTCTGCAACCCCTGCTGCAAGAG 960
QY 1003 CCGATGCTTAAAGCCCAACCCGATCCCATCTCTGAGAGCTGCGCAAGTGTCTGAT 1062
DB 961 CCGATGCTTAAAGCCCAACCCGATCCCATCTCTGAGAGCTGCGCAAGTGTCTGAT 1020
QY 1063 GGACTGGGGCTGAGCAGAGCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122
DB 1021 GGACTGGGGCTGAGCAGAGCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1123 TAG 1125

PHLLPSPPPASDLSPAVAPATRLGPIILLERBQSGSTRALHCPRTBTCKRYPAS
 EAQAVLAFAVRLPTHQVAPRTVLLSRLVLTPTKTHDLSHVSRRGTPSBEAA
 GLPRQMASAVAHCHKQLVRLDLRLRFVNSCBRTLVLENEEDACVPMNSDLSLD
 THACPAPVGBEILSSRPSYSGKADYSLVALPTMAGVYPPHDSPPVLLFGKIRG
 TPLPGLSAPARCLIRCLRKPSBRVLVALILHPLMLBHDGRVSPQSDRREMDQ
 VVDDGRLSEAESEGVLYG"

BASE COUNT 231 a 390 c 385 g 249 t 1 others
 ORIGIN
 Query Match 32.1%; Score 661.4; DB 10; Length 1256;
 Best Local Similarity 75.7%; Pred. No. 4,8e-117;
 Matches 818; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 17 GCGCCGCGCCCAACGCGGAAAGAGATGCAAGCCCTCTGCTGCTCTG 76
 DB 150 GCGCGCAGCGGCTCTAGAGCTCCAGAGCAAGATGCGAGCTACACTCTGCTCTCTG 209
 QY 77 CGGGTTCCTGTCAGAGAAAGAGGCTTGAAGTGAACAATTGATTAACGAGGCTC 136
 DB 210 CTGATGTTCTCGAGAGAAAGACCTTGAAGTTGAACAATATGATGACCAAGTGT 269
 QY 137 CCGTCCAGAAACGAGCTCGAAGTGGGCGCCAGCCAGACTGCCCCCTGCTGTGCCCC 196
 DB 270 CAGTCCAAAACGAGTGAAGATGAGCTGAGCCGAGCACTCCAGGCTGCTGCCCC 329
 QY 197 TGAAGCCACTACTGCTCCAGATGTCGCAACTGCTGTGGCACTGCTCTCCGCTTTGGGC 256
 DB 330 CCAAGCCACTCTCCGCTCAGACTTGTACCTGTGTGGGCCCCCTGCAACTGACTGGGGC 389
 QY 257 CCTATGCTCTCTGGAAGCCGAGAGGGGCGGGGCTACCGGGGCTGCACTGACCTTA 316
 DB 390 CTATATCTCTTTTGAACGAGAGCAGAGCTGAGCTATCGAGCCCTGCACTGCCCCA 449
 QY 317 CAGGCACTGAGTATACCTGCAAGGTATACCCGCTCAGAAAGCCCTGAGCGTGTGAGC 376
 DB 450 CAGGCAAGAGTACACTGCGAAGGTATACCTGCAAGGAGCCAGGCGGTGTGCGAC 509
 QY 377 CTAAGCGCGGCTGCCCCGCAAGATGTGCTCGGCCCACTGAAGTCTGTGCTGTA 436
 DB 510 CTATGCGCGGCTGCTTACCAAGATGTGCGCTCCCAAGAGTCTGTGCGGCT 569
 QY 437 CCGAGCTCTCTAAGCTTTTCACTCGAACCCATGGAGGACATGCAAGCTGTGTGGA 496
 DB 570 CTGAGCTCTTACATCTTTTCAAGAAACCAATGGAGACTTGCAGAGCTGTGTGCGCA 629
 QY 497 GCGGCAACGATCTCTGAGCTGAGGCTGCGCTCTTCCGCAAGTGGCCACGCGCC 556
 DB 630 GCGGCGCGGATATCCAGAGTCCGAGGCTGCGGGCTCTTCCGCAAGTGGCTAGTCCG 689
 QY 557 TGGCGCACTGCAACGAGCTGTGCTCGCTGCTGCTCAAGCTGTGTGCTTTGTCT 616
 DB 690 TGGCACTGCAACGAGCTGTGCTGTGCTGCGCACTCAAGCTGTGTGCTTTGTCT 749
 QY 617 TCGCTGACCTGAGAGAAAGCTGTGCTGAGAACTGAGAGACTCTGCTGCTGTA 676
 DB 750 TCAAGCACTGTGAGAGGACGAACTGTGCTGAGAACTGAGAGATGCTGCTGTA 809
 QY 677 CTGAGCGAGATGATCTCTGTGAGCAAGCAACGCTGCCAGCTCACTGTGAGACTGAGA 736
 DB 810 CTGAGATGAGATGATCTCTGTGAGCAACGATGCTGCCCTCACTGTGAGCACTGAGA 869
 QY 737 TACTCAGCTGACGAGGCTCTGATCTCGGCAAGGAGCGATGTCTGAGGCTGTGGCGTGG 796
 DB 870 TACTCAGCTGACGAGGCTCTGATCTGTGCAAGGAGCGATGTCTGAGGCTGTGGCGTGG 929
 QY 797 CGCTCTTCAACATGCTGTGAGCGGCACTACCTTTCCAGAACTCGAGACTGTCTGCTCT 856
 DB 930 CGCTCTTCAACATGCTGTGAGCGGCACTACCTTTCCAGAACTGTGAGCGTGTGCTCT 989
 QY 857 TCGGCAAGATCGCGCGGCGGCTTACGCTGTGCTGAGAGCTCTCGGCGCTGTGCGGCT 916
 DB 990 TTGGCAAGATCGGTANAGGACCTTGTGCTGTGAGGCTCTATCAGGCCCAAGCCGCT 1049

QY 917 GTCTGCTGCTGCTGCTCTCTCTGTCGAGAGCCAGCTGAACGGCTCAACAGCCACAGCACTCC 976
 DB 1050 GTCTGATCCGCTCTCTCTCTCCGAGAGAACTTCAAGGCACTTGTGCGCTGTGAAATCC 1109
 QY 977 TCGTGAACCCCTGAGCTGCGACAGAGACCCGATGCTTAGCCCCCAACCCGATCCCATCTCT 1036
 DB 1110 TCTTGATCTCTCTGTTGAGAGAGATCAAGGCGGAGCTCTCTCTCCACAGCTGAACCGAA 1169
 QY 1037 GAGAGGCTGCGCAGATGCTCTGATGAGACTGAGGCTGAGCAAGAGCAAGAGAGAGAG 1096
 DB 1170 GAGAGATGAGCAAGATGCTCTCCAGATGAGCCACAGCTGAGAGAGCTGAGAGAGAGAG 1229

RESULT 12
 AX364906 1076 bp DNA linear PAT 15-FEB-2002
 LOCUS
 DEFINITION Sequence 57 from Patent WO0206315.
 ACCESSION AX364906
 VERSION AX364906.1 GI:18696795
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Bacteria; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCES
 1 Mintz, L., Freilich, S. and Bernstein, J.
 Novel nucleic acid and amino acid sequences
 Patent: WO 0206315-A 57 24-JAN-2002;
 Compugen Ltd. (IL)
 location/Qualifiers

FEATURES

source
 1..1076
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 172 a 352 c 346 g 194 t 12 others

Query Match 31.3%; Score 645.2; DB 6; Length 1076;
 Best Local Similarity 98.2%; Pred. No. 6.2e-114;
 Matches 670; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

QY 1 GCTGTGAGCCCGGCGGCGCGCGGCGCCAGCGGAAAGAGGCGGAGATGCGAGCCACC 60
 DB 281 GCTGTGAGCCCGGCGGCGCGCGGCGCCAGCGGAAAGAGGCGGAGATGCGAGCCACC 340
 QY 61 CCTGTGAGCTCTCTGCGGCTTCTGTCCAGAGAAAGAGGCTTGAATGATGACAAC 120
 DB 341 CCTGTGAGCTCTCTGCGGCTTCTGTCCAGAGAAAGAGGCTTGAATGATGACAAC 400
 QY 121 TTAGATACGAGAGGCTCTGCTGCAAGAAAGAGCTGAAAGTGGGCGCCAGCCAGCTGCC 180
 DB 401 TTAGATACGAGAGGCTCTGCTGCAAGAAAGAGCTGAAAGTGGGCGCCAGCCAGCTGCC 460
 QY 181 CCGTGCCTGTGAGCCGCGGAGCCAGCACTGCTCAAGATGATGATGCACTGTGTGGCACT 240
 DB 461 CCGTGCCTGTGAGCCGCGGAGCCAGCACTGCTCAAGATGATGATGCACTGTGTGGCACT 520
 QY 241 GCTCTCCGCTTGTGAGGCTTATGTCTCTGAGAGCCGAGAGAGGCGGCGGCTCAACCG 300
 DB 521 GCTCTCCGCTTGTGAGGCTTATGTCTCTGAGAGCCGAGAGAGGCGGCGGCTCAACCG 580
 QY 301 GCGCTGACCTGCTTCAAGGCACTGAGTATCTGCAAGATGATACCCGCTGCAAGAAAGCC 360
 DB 581 GCGCTGACCTGCTTCAAGGCACTGAGTATCTGCAAGATGATACCCGCTGCAAGAAAGCC 640
 QY 361 CTGAGCGTGTGAGAGCTTACAGCGCGGCTGCGCCGAGCAAGAGATGATGATGCGGCACT 420
 DB 641 CTGAGCGTGTGAGAGCTTACAGCGCGGCTGCGCCGAGCAAGAGATGATGATGCGGCACT 700
 QY 421 GAGGCTGTGAGTGAACCAAGCTCTCTCAAGCTTTTCACTGAGACCAATGGAGGACATG 480
 DB 701 GAGGCTGTGAGTGAACCAAGCTCTCTCAAGCTTTTCACTGAGACCAATGGAGGACATG 760

OY 481 AACAGGCTGAGGCGAAAGGCGGACAGGATATCCCTGAGGCTGAGAGGCTGACGGTGTCTTCCG 540
 Db 761 CACAGCCTGAGTGCAGAGGCSGCCACCGATATCCCTGAGGCTGAGAGGCTGACGGTGTCTTCCG 820
 OY 541 CAGATGGCCACCGGCCCTGGCGCAGCACTGTACACGACCGGTCTTGTCTCTGCTGATCTCAAG 600
 Db 821 CAGATGGCCACCGGCCCTGGCGCAGCTGTACACGAGCAGCGGTCTTGTCTCTGCTGATCTCAAG 880
 OY 601 CTGTGTCCGCTTGTGTCTTGGCTGACCGTGAAGAGAAAGAAAGCTGTGTCTGTGAAGAACTTGAG 660
 Db 881 CTGTGTCCGCTTGTGTCTTGGCTGACCGTGAAGAGG-AGAAAGCTGTGTCTGTGAAGAA--CTTGA 937
 OY 661 GACTCTGAGTGTGAGTCTGAGTGGG 682
 Db 938 GACTCTGTGCTGTACTGTGGG 959

LOCUS	AX364921	1076 bp	DNA	linear	PAT 15-FEB-2002
DEFINITION	Sequence 72 from Patent WO0206315.				
ACCESSION	AX364921				
VERSION	AX364921.1	GI:18696810			
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 Mintz, L., Freilich S. and Bernstein, J.				
AUTHORS	Novel nucleic acid and amino acid sequences				
TITLE	Patent: WO 0206315-A 72 24-JAN-2002;				
JOURNAL					

FEATURES	source	Location/Qualifiers
		1..1076
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:3606"
BASE COUNT	172 a	352 c 346 g 194 t
ORIGIN		12 others

Query Match	Similarity	31.3%	Score 645.2	DB 6	Length 1076
Best Local	Similarity	96.2%	Pred. No. 6,2e-114		
Matches	670	Conservative	3	Mismatches	6
				Indels	3
				Gaps	2
Qy	1	GCTCTGAGCCCCCGGCGGCCCGCGGCCCGACGCGAGCGGGCGGAGTTCGAGCCACC	60		
Db	281	GCTCTGAGCCCCCGGCGGCCCGCGGCCCGACCGGAGACAGCGGGCGGAGTTCGAGCCACC	340		
Qy	61	CCTCTGGGTGCTCCCTGCGGGGTTCCCTGTGTCAGAGAAAGCGGTTGAGTTGAGTGAAC	120		
Db	341	CCTCTGGGTGCTCCCTGCGGGGTTCCCTGTGTCAGAGAAAGCGGTTGAGTTGAGTGAAC	400		
Qy	121	TTAGATACCGAGCGTCCCGTCAGAAAGAGCTCGAAGTGGGCCCGACCGACACTGCC	180		
Db	401	TTAATATACCGAGCGTCCCGTCAGAAAGAGCTGAAATGGGCCCGACCGACACTGCC	460		
Qy	181	CCCTGCTGTGTGCCCTCTAGAGCCCACTACTGCTCCAGTCTGTGCACTGTGTGGCACT	240		
Db	461	CCCTGCTGTGTGCCCTCTAGAGCCCACTACTCTCCAGTCTGTGCAACTGTGTGGCACT	520		
Qy	241	GCCTCCCGGCTCTTGGGGCCCTATGTCTCTCTGAGGCCGAGGAGGGCGGGCCCTACCGG	300		
Db	521	GCCTCCCGGCTCTTGGGGCCCTATGTCTCTCTGAGGCCGAGGAGGGCGGGCCCTACCGG	580		
Qy	301	GCCTGCACTGCGCTTACAGGCACTGAGTATACCTGCAAGGTGTACCCGTCAGAGAGCC	360		
Db	581	GCCTGCACTGCGCTTACAGGCACTGAGTATACCTGCAAGGTGTACCCGTCAGAGAGCC	640		
Qy	361	CTGGCCGTGCTGAGAGCCCTTAGCGCGCGTGGCCCCCGACAGCACTGTGTGCTGGCCCACT	420		
Db	641	CTGGCCGTGCTGAGAGCCCTTAGCGCGCGTGGCCCCCGACAGCACTGTGTGCTGGCCCACT	700		

[illegible]

LOCUS	AB020967	2004 bp	mRNA	linear	ROD 14-MAY-1999
DEFINITION	Rattus sp. mRNA for kinase, complete cds.				
ACCESSION	AB020967				
VERSION	AB020967.1 GI:4827158				
KEYWORDS	kinase; NIPK.				
SOURCE	Rattus sp.				
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases)				
AUTHORS	Mayumi-Matsuda,K., Kojima,S., Suzuki,H. and Sakata,T.				
TITLE	Identification of a novel kinase-like gene induced during neuronal cell death				
JOURNAL	Biochem. Biophys. Res. Commun. 258 (2), 260-264 (1999)				
MEDLINE	99262087				
PUBMED	10329375				
REFERENCE	2 (bases 1 to 2004)				
AUTHORS	Matsuda,K.M., Kojima,S. and Nakayama,T.				
TITLE	Direct Submission				
JOURNAL	Submitted (09-DEC-1998) Keiko Mayumi Matsuda, SHIONOGI and Co. Ltd., Shionogi Institute for Medical Science, 2-5-1 Mieshima, Settsu-shi, Osaka 566-0022, Japan				
	(E-mail:keiko.matsuda@shionogi.co.jp, Tel:81-6-6382-2612(ex.465), Fax:81-6-6382-2598)				

```

FEATURES
Source
Location/Qualifiers
1..2004
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="taxon:10118"
/cell_line="PC12"
/cell_type="pheochromocytoma"
1..2004
/gene="NIPK"
154..1203
/gene="NIPK"
/note="neuronal cell death inducible putative kinase (NIPK); induced by NGF-depletion"
/product="kinase"
/codon_start=1
/protein_id="BAA7582.1"
/db_xref="GI:4827159"
/translation="MRATSLAASADVPCKRKPLSPDDNINVECVLKRVRDEPFCPT
PSLPSPDISPAVAPAPATRLGPYILLREQNCYRALHCPGTGTYTCVYPASAOAV
LAYAPALPTHOHVAAPTEVILGSQLYTFPTKTHGDIHSVRSRGIPEPEAALFPQ
MAAVVHCHHGGLILDKLRVPSFCERTKLISLBDACVWGTGDSLIMDGACFP
AAYGPEILSRPSYSGRADVWSLGVALPTMLAGRYVPROSPBALPGKIRGFLAP
EGASASARCIRCLLRBPSESRVLVAGLILPWLREDCQVSPSPSRDMQVVPDGG
PQLEAEKEGSGVLYG"
2004
polya_site

```

polyA_site

2004

```

/gene="NIPK"
/note="27 a nucleotide"
BASE COUNT      410 a      563 c      590 g      441 t
ORIGIN
Query Match      30.4%; Score 626.4; DB 10; Length 2004;
Blast Local Similarity 72.1%; Pred. No. 2.9e-110;
Matches 883; Conservative 0; Mismatches 311; Indels 30; Gaps 4;

17  GCGCCGCGGCGCCAGCGGGAAGAGATGCGAGCCGCTCTGCTGCTCTG 76
122 GGGGCGAGGGGGCTCTAGACCTCAGAGCAAGATGCGAGCCATCTCTGCTCTG 181
77  GGGGTTCCCTGTCAGAGAGAGGGTTGAGTTGATGACACTTGAATCCAGGCTC 136
182 CCGATGTTCCCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 241
137 CCGGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196
242 CAGCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288
197 TGAGCCCACTACTGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 256
289 --AGCCTACTCTCCGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 346
257 CCTATGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
347 CCTATATCTCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 406
317 CAGGACATGAGATATACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 376
407 CAGGACATGAGATATACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 466
377 CTAAGCGCGCGCTGCGCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
467 CGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 526
437 CCCAGCTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
527 CTGAGCTCTATATACCTTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 586
497 GCGGCAACCGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556
587 GCGGCGCGGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 646
557 TGCGGCACTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
647 TGCGCACTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
617 TCGCTGACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 676
707 TCAAGCACTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 766
677 CTGGGCAAGATGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736
767 CTGGACCAAGAGAGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
737 TACTGAGCTCAGCGGCTCTATCTGCGGCAAGAGAGAGAGAGAGAGAGAGAG 796
827 TACTGAGCTCAGCGGCTCTATCTCTGCGGCAAGAGAGAGAGAGAGAGAGAG 886
797 GCGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 856
887 GCGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
857 TCGGCAAGATCGCGCGGCGCTCTATCTGCGGCAAGAGAGAGAGAGAGAGAG 916
947 TTGGCAAGATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
917 GTCTGAGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 976
1007 GCGCTCATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1066

```

```

Qy 977 TCCGCAACCCCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1036
Db 1067 TGCTACATCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126
Qy 1037 GAGAGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1096
Db 1127 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
Qy 1097 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1155
Db 1186 -----GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1234
Qy 1156 AGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
Db 1235 TGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1292
Qy 1216 TCAGTGCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1239
Db 1293 TAACTGCTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1316

```

```

RESULT 15
BC012955      1969 bp. mRNA. 11near. ROD 16-APR-2003
LOCUS         Mus musculus induced in fatty liver dystrophy 2, mRNA (cdna clone
DEFINITION    MGC:18731 IMAGE:3980838), complete cds.
ACCESSION     BC012955
VERSION       BC012955.1 GI:15277944
KEYWORDS      MGC.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus

```

```

REFERENCE
AUTHORS       Krausner,R.D., Collins,P.S., Wagner,L., Shermen,C.M., Schuler,G.D.,
              1 (bases 1 to 1969)
              Xlauner,R.D., Collins,P.S., Wagner,L., Shermen,C.M., Schuler,G.D.,
              Altschul,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhac,N.K.,
              Hopkings,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
              Diatchenko,L., Marubina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
              Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
              Scheetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
              Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
              Abramson,R.D., Mulhaly,S.J., Bosak,S.A., McMan,P.J.,
              McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
              Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huily,S.W.,
              Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
              Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shvchenko,Y.,
              Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvchenko,Y.,
              Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,B.D.,
              Dickson,M.C., Rodriguez,A.C., Gilmwood,J., Schmutz,D., Myers,R.M.,
              Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smallov,D.B.,
              Schnerch,A., Schein,J.B., Jones,S.J., and Mair,M.A.
              Generation and initial analysis of more than 15,000 full-length
              human and mouse cDNA sequences
              Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```

```

TITLE
JOURNAL       JOURNAL
MEDLINE       MEDLINE
PUBMED        22388257
REFERENCE     12477932
AUTHORS       Strausberg,R.
TITLES        Direct Submission
JOURNAL       Submitted (20-APR-2001) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA

```

```

REMARK
COMMENT       NIH-MGC Project URL: http://mgc.ncl.nih.gov
              Contact: MGC help desk
              Email: gcgaps-remail.nih.gov
              Tissue Procurement: Gilbert Smith, Ph.D.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)
              Sequencing Center
              Center code: BCM-HGSC

```


Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
 Kowib, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Namballi,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
 Series: IRAC Plate: 23 Row: k Column: 9
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, similarity but not identity to protein.

FEATURES

source

1. 1969
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:18731 IMAGE:3980838"
 /issue_type="Mammary tumor. Metastatic origin. TGF alpha
 model. 10 month old virgin mouse. Taken by biopsy."
 /clone_1lb="NCI CGAP_Mam1"
 /lab_host="DHI0B"
 /note="Vector: pCMV-SPORT6"
 1. 1969
 /gene="Ifid2"
 /note="synonym: TRB-3"
 /db_xref="LOCUSID:228775"
 /db_xref="MGI:1345675"
 177. 1163
 /codon_start=1
 /product="TRB-3"
 /protein_id="AAH12955.1"
 /db_xref="GI:15277945"
 /db_xref="LOCUSID:228775"
 /translation="MRATPLAASADVSCRKPKLPFDNDIDAKPVLKRVBDEPGL
 PSLLPSPSPASDLSPPAVAPARLIGPILREOGSCSYALHCPETBYCKVPPAS
 BAQAVLAPVRLPTHOVARPTLVLSGLLYLFTTHGDLILSVRSRGIPEPPAS
 KPRQMAVAHCHGILVLDLKLRFVSNCRRTLVLENBDACVWMSDSDSLND
 KACPRVAVGEILSRPSYSGPVLKFKIRGTFRALPBGISARCLIRLKLKRSBS
 RLVALGILHPLHREDHGRVSPQSDREMDVVPDQPLEABEGVGLYG"

CDS

BASE COUNT 413 a 560 c 572 g 424 t
 ORIGIN

Query Match 26.64; Score 546.8; DB 10; Length 1969;

Best Local Similarity 61.94; Pred. No. 6.1e-95; Indels 197; Gaps 16;

Matches 1247; Conservative 0; Mismatches 572; Indels 197; Gaps 16;

17 GCGCCCGGCGCCAGCGGGAAGAGCGGCGAGATGCGAGCCACCCCTGCTGCTCTG 76
 145 GGGCGCAGGCGGCTGAGGCTCCAGGAGCAAGATGCGAGCTACACCTGCTGCTCTG 204
 77 CGGATTCCTCTGTCAGAGAGAGCGGTTGAGTTGATGCACTTAGATACGAGCGTC 136
 205 CTGATGTTTCTGCGAGAGAGAAACGTTGAGTTGATGCAATGATGCGCAAGTGC 264
 137 CGGTCCAGAAAGAGCTCGAAGTGGGCGCCAGCCAGACCTGCCCCCTGCTGTTGCCC 196
 265 CAGTCTTAAACAGATGAGATGAGCTGAGCCCGGACCACTCCAGCTCTGCCCC 324
 197 TGAGCCACCTACTGCTCAGATGTCGCACTGCTGAGCCACTGCTCCCTCTTGGAG 256
 325 CCAGCCACCTCCCGCTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384
 257 CCTATGCTCTCTGCGAGAGCGGAGAGGCGGCGGCTTACCGGCGCTTGCATGCTG 316
 385 CTATATATCTTTTGAACAGAGAGAGAGCGCTGAGCTTATGAGCCCTGCTGCTG 444
 317 CAGGCACTAGTATACCTGCAAGGTGTACCCCTGTCAGAGAGCCCTGAGCTGAGAG 376
 445 CAGGCAAGAGTACCTGCAAGGTGTACCTGCTGCGAGAGCCCGAGGCGGTGCTGAG 504
 377 CTTAGCGGCGGCTGCCCCGCGCAGACAGATGTGCTGCGGCGCACTGAGGTCTGCTG 436

DB 505 CTTATGCGCGGCTGCTTACCAACAGATGTGCGCTCCACAGAGGTCTGCTGCGCT 564
 437 CCCAGCTCTCTACGCTTTTTCATCTGGAACCAATGAGGAGATGCAAGCTGTGCGAA 496
 565 CTCGCTCTCTTACATCTTTTTCAGGAAGCCATGAGGAGCTTGCACAGCTGTGCGCA 624
 497 GCGGCAACGATTCCTGAGAGCTGAGGCTGCGCTCTTCCGCAATGAGCCACCGCC 556
 625 GCGCGCGGATTCCTGAGAGCCGAGGCTGCGCTCTTCCGCAATGAGCTGTGCG 684
 557 TGCGCACTGTACACAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
 685 TGCGCACTGTACACAGACAGGCTTGTCTTGGCGGACTTCAAGCTGTGCTGCTGCT 744
 617 TCCTGACCTGTGAGAGAGAGAGTGTGCTGAGAACTTGAAGACTCTGCTGCTG 676
 745 TCAGCACTGTGAGAGAGAGAGTGTGCTGAGAACTTGAAGACTCTGCTGCTGAG 804
 677 CTGGGCGAGATGATTCCTGCTGAGAGAGAGAGCTGAGAGCTTACCTGAGAGAG 736
 805 CTGATGATGATGATCTCTGTGAGAGAGAGAGAGTGTGCTGCTGCTGAGAGAG 864
 737 TACTCAGCTACAGGCTCTCATTAATCGGCAAGAGAGAGAGTGTGAGAGCTGAGAG 796
 865 TACTCAGCTACAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 885
 797 CGCTCTTACAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 856
 886 -----ACTGAG 906
 857 TCAGCAAGATTCGCGCGGCGGCTCTAGCGCTTGTGCTGAGAGCTTGTGCTGCTG 916
 907 TTGCAAGATTCGAG 966
 917 GTCTGATTCGCTGCT 976
 967 GTCTGATTCGCTGCT 1026
 977 TCCTGACAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1036
 1027 TCTTGATTCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086
 1037 GGAAGAGCTGCGAG 1096
 1087 GGAAG 1145
 1097 GAGACAG 1156
 1146 -----GTGAG 1194
 1157 GTGATGATGAG 1216
 1195 GTGATGATGAG 1252
 1217 CAGTGTCTTCCAG 1276
 1253 AAGTGTCTTCCAG 1282
 1277 GCTTGTCTTCCAG 1336
 1283 CGTGTCTTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1342
 1337 TCCTGCTGAG 1396
 1343 TAGGTGTGTGAG 1402
 1397 AAAGTGTCTTCCAG 1456
 1403 AAGTGTCTTCCAG 1439
 1457 GTACATTTGTACAG 1516

Db	1440	GCAGTTTCCATACC-GGGGCTGTCTTCTCTGTGTGTGT-----TCATGG	1483
Oy	1517	GACAAATCCCTTTCACAAACCAACGCTGCCCTTTGTATCTGTATACCTTTTCAGAAAGG	1576
Db	1484	GACAGTCCCTTTTATCACTGCCCCCAGCTGCC-TTCTGTCTGTGCTCTTTAA---CAAG	1539
Oy	1577	GAGGTATCCCTGTGCAAAAGGCTCAGGCTCTC---CCCTGCAATCAGAGCCCAAGCC	1633
Db	1540	GAACATCTCTGTGCAAAAGACAAAGTCTCTCATCATCATGTGCTCAGATCCAAAT	1599
Oy	1634	CAGCTCACTCTGGGAACTGTGTCTCCCAACATCTGTCTCTCTGTATTAAGATTTCTCT	1693
Db	1600	CAGCGCGGGCTGGGATTTGTGGCTCTCAAGCAGAGAACTGTTTGAAAG-----	1648
Oy	1694	TCCAGCGCTTAAGCCTGGGATTTGGGCGCAGAGATTAAGATTCAAACTATAGAGCTAGTCT	1753
Db	1649	---AGACTTGGCTGTGGGATTCAGAGCCAAAGAGAACTCAATTAAGAGGACATT--	1703
Oy	1754	TGTCTAATCTCAAGACTGTTCTTGGAATGAGGGTCCAGGCTGTCAACCATGG-GGCTTCTG	1812
Db	1704	----GACTCAAGGAATAGTCTTGGAATAGGATTCGGGCTGTGTATTAACCTGTGTACCTTCG	1759
Oy	1813	ACCTGAGCACAAGGTTG-AGGGACAGGATTAAGCAGAGGCTGTGTCCTGTGSCCACTGGA	1871
Db	1760	GCTGAGGACACCAAAATTGAGGGGGAGAACCAAGGAGGGTCTGTCTTACTGCTCTACAG--	1817
Oy	1872	AAGTCCCAAGTGGGACTCTTCTGGGSAACATTGGGGGTCCAAATCCCAAGTCCATACTCT	1931
Db	1818	----ACAAATCTCAGGCCCTGTGTTTGGGACCTTTAGTACCCAGGCCCAAGTCAGTTTGA	1873
Oy	1932	AGGTTTGGATACCATGAGTATGATGTGTTA-CTGTGCTTAATTAAGAGAAATTATGAA	1990
Db	1874	TGACTTTGTGTGTTAAGTGTGTATTTTCATCATGTGTCCTCAATTAAGAGAGCTGTGTG	1933
Oy	1991	ATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2026
Db	1934	GTTTGATATGTGAATAAAAAAAAAAAAAAAAAAAAAA	1969

Search completed: January 16, 2004, 14:49:00
Job time : 7734.16 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:55:46 / Search time 41 Seconds
(without alignments)
2253.240 Million cell updates/sec

Title: US-09-909-474D-2

Perfect score: 1891
Sequence: 1 MRAATPLAAPAGSLSRKKRLR.....GLGLDPAAREEGDREVLVYG 358

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	756	40.0	343	11 Q8K017	Q8K017 mus musculus
2	754.5	39.9	343	6 Q28283	Q28283 canis fam1
3	753.5	39.8	343	4 Q92519	Q92519 homo sapien
4	753	39.8	343	11 Q8K4K3	Q8K4K3 mus musculus
5	749.5	39.6	364	11 Q9E0L6	Q9E0L6 rattus norv
6	744.5	39.4	372	4 Q86R08	Q86R08 homo sapien
7	742.5	39.3	372	4 Q9H2I8	Q9H2I8 homo sapien
8	742.5	39.3	372	11 Q91W04	Q91W04 mus musculus
9	729.5	38.6	372	11 Q8K4K4	Q8K4K4 mus musculus
10	653	34.5	257	11 Q8R2V8	Q8R2V8 mus musculus
11	611.5	32.3	424	4 Q15180	Q15180 homo sapien
12	435	23.0	484	5 Q9V321	Q9V321 drosophila
13	326	17.2	218	11 Q8B8S7	Q8B8S7 mus musculus
14	326	17.2	379	11 Q8B3R9	Q8B3R9 mus musculus
15	326	17.2	388	11 Q8B2X3	Q8B2X3 mus musculus
16	315	16.7	461	10 Q9LGV5	Q9LGV5 oryza sativ

17	314.5	16.6	706	3 Q8J2N0	Q8J2N0 fusarium ox
18	305	16.2	474	10 Q946T7	Q946T7 oryza sativ
19	302.5	16.0	646	4 Q96CV1	Q96CV1 homo sapien
20	299	15.8	914	5 Q19469	Q19469 caenorhabdi
21	297.5	15.7	446	10 Q8H5S1	Q8H5S1 oryza sativ
22	297.5	15.7	778	4 Q8TDC2	Q8TDC2 homo sapien
23	297	15.7	504	11 Q8K0J7	Q8K0J7 mus musculus
24	297	15.7	880	3 Q9V880	Q9V880 cochlilobol
25	296	15.7	671	3 Q96W17	Q96W17 trichoderma
26	295	15.6	794	4 Q8TDC3	Q8TDC3 homo sapien
27	292.5	15.5	651	13 Q91821	Q91821 xenopus lae
28	291	15.4	602	3 P87209	P87209 kluyveromyc
29	289.5	15.3	688	5 Q61298	Q61298 halocynthia
30	289	15.3	688	5 Q95UP4	Q95UP4 ancylostoma
31	288	15.2	477	10 Q9LW6	Q9LW6 oryza sativ
32	285	15.1	514	10 Q9W726	Q9W726 lycopersico
33	284	15.0	504	10 P93113	P93113 cucumis sat
34	284	15.0	512	10 P92958	P92958 arabidopsis
35	283.5	15.0	643	11 Q61804	Q61804 mus musculus
36	283.5	15.0	643	11 Q61846	Q61846 mus musculus
37	283.5	15.0	651	4 Q14680	Q14680 homo sapien
38	283	15.0	458	10 Q8LK24	Q8LK24 glycine max
39	283	15.0	512	10 P92968	P92968 arabidopsis
40	282	14.9	444	10 Q9FUK2	Q9FUK2 arabidopsis
41	281.5	14.9	419	4 Q9H7H6	Q9H7H6 homo sapien
42	281	14.9	535	10 Q8RMD2	Q8RMD2 arabidopsis
43	280	14.8	339	10 Q39868	Q39868 glycine max
44	280	14.8	444	10 Q8RWC9	Q8RWC9 arabidopsis
45	280	14.8	480	10 Q9LUP6	Q9LUP6 arabidopsis

ALIGNMENTS

RESULT 1
ID Q8K017 PRELIMINARY; PRT; 343 AA.
AC Q8K017;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE TRB-2 (CSFM ORF protein homology).
OS Mus musculus (Mouse).
OC Bakaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Cerebellum, and Retina;
RX MEDLINE=2354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; BC034338; AAC34338.1; -;
DR EMBL; AK044747; BAC32063.1; -;
DR EMBL; AK080064; BAC37820.1; -;
DR EMBL; AK082329; BAC38467.1; -;
DR InterPro; IPRO00719; Prot kinase.
DR InterPro; IPRO02290; Ser Thr kinase.
DR InterPro; IPRO02245; Tyr kinase.
DR pfam; PF00069; kinase; 2.
DR PRODOM; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.

Seq	SEQUENCE	343 AA;	38772 MW;	941887/AC19FCC23F CRC64;
Query Match	40.0%;	Score 756;	DB 11;	Length 343;
Best Local Similarity	46.9%;	Pred. No. 1.8e-55;		
Matches 164;	Conservative 48;	Mismatches 108;	Indels 30;	Gaps 6;
Qy	2	RATPLAAGSLSRKGRLELDNDLPTERPVOKRARGQPRLPCLPLSPPTAPDRATY	61	
Db	5	RSTPTTIRYGRSRRKTKDDPER-LSSISAR-----PSGRSPILSGSPSPERTNLSHC	57	
Qy	62	VATASRLGPRVYLLEBEGGGRVYALHCSPTGYTCKVYV--GEALVLEPYARLPKH	118	
Db	58	V---SCIKYLLLEPLEGDHVFRAVHLHSGEELVCKVSEISCYGSLA---PCFLCSHS	111	
Qy	119	HVARTEVLATQGLYAFPTTHGMSHLYNSRHHIPPEPAVLPRQMATLHCHQGL	178	
Db	112	NINQTELLLESTKAVVEFENSYGMHSFVRCCKLRBEZARLYQTASAVAHCHDGL	171	
Qy	179	VLRDLKCRPFVADREKRKLVLENLEDSCVLTGDPDSDMDKXACPAVYGPBILSRASYS	238	
Db	172	VLRDLKAKKPFKDEBTRRVKLESLEDAYILRGDDSDLSKDKGCPAVYSPBILNTSGYS	231	
Qy	239	GKADVMSLGVALLPMLAGHPFPOSEPVLLFGKTRRCQAYVALPALSAPARCLVCLLR	298	
Db	232	GKADVMSLGVALLPMLAGHPFPHIDIPSSLSKTRRQCPNPELSPAKKLIRSLRR	291	
Qy	299	EPARLELTGILLHPMLRODPMPLAPTRSHLMEAAQVVPDGLDEAREE	348	
Db	292	EPSERLTSGEILDHMPSTD-----PSVNSGSGAEACDQ	327	
RESULT 2				
Q28283		PRELIMINARY;	PRT;	343 AA.
ID	Q28283			
AC	Q28283			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	C5FW ORF protein.			
GN	C5FW ORF.			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
CK	NCBI_TaxId=9615;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Thyroid;			
RX	MEDLINE=97067069; PubMed=8910471;			
RA	Wilkin F., Savont V., Radulescu A., Petermans J., Dumont J.B.,			
PA	Maenhaut C.,			
RT	"Identification and Characterization of Novel Genes Modulated in the			
RT	Thyroid of Dogs Treated with Methimazole and Propylthiouracil."			
RL	J. Biol. Chem. 271:28451-28457(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Thyroid;			
RX	MEDLINE=98000062; PubMed=9342215;			
RA	Wilkin F., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,			
PA	Dumont J.B., Maenhaut C.,			
RT	"Characterization of a phosphoprotein whose mRNA is regulated by the			
RT	mitogenic pathways in dog thyroid cells."			
RL	Eur. J. Biochem. 248:660-669(1997).			
DR	BMJL; X99144; CAA67581.1; -.			
DR	HS8F; O63450; 1A06.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	Pfam; PF00069; pkinase; 1.			
DR	ProDom; PDD00001; Prot_kinase; 1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
DR	ATP-binding; Transferase.			
DR	SEQUENCE 343 AA; 38786 MW; BF6D1300DACB84FA CRC64;			
Query Match	39.9%;	Score 754.5;	DB 6;	Length 343;
Best Local Similarity	47.3%;	Pred. No. 2.4e-55;		

	Matches	167,	Conservative	50;	Mismatches	115;	Indels	21;	Gaps	6
Oy	2	RATPLAAPAGSLSSKKKLELDNDLDTERRPVQKARSGOPRLPCLLPSPPTADRA	TA 61							
Dd	5	RSTPTIARVGRSHNKTDFFE--LSIRSAB-----PSQSFPMLGSPSPETPNLSHC	57							
Oy	62	VATASRLGPYLLAPBEGGARAYRALHCTGTGYTCCKVPV---QBALVLBPYALPPHK	118							
Dd	58	V---SCIGKYLLPLEBEDHVFRVHLHSGBELVCVKPFDISCYESLA---PCFLSAS	111							
Oy	119	HVARPTVIAGTOLLVAFFTTTHDDMSLVSRHRIPERBAVLPRONATALAHCHQGL	178							
Dd	112	NINOTTEILIGETAAYYPFERSYGDMHSFATCCKLRBBEARLYOJASNAHCHDGL	171							
Oy	179	VLRLDKLCRFVAFADRRKRKLVLNMLEDSCVLTGPDSDLMDXACPAVYGGPIILSRASYS	238							
Dd	172	VLRLDKRKRFKFOBERRTVLESLEDAYLIRGDODDSLXKHGCCPAYSPRILTNSGSYS	231							
Oy	239	GKAADWSLGVALTMTLAGHYPPQDSEVULLFGKTRRCAVALPAGLSAPACLVRCLR	298							
Dd	232	GKAADWSLGVALTMTLVGRYPFDHIEPSSLFSKIRQCQFNIPELTSKACCLRSILRR	291							
Oy	299	EPARBLTAGTILLHPWLRODPMPLAPTSSHLMBAOVVPGDGLDEAREEBCD	351							
Dd	292	EPERLISQBIILDPHFSTD-----FSVSNSGCGAKESVDQLVPDYMNENLD	339							
 RESULT 3 092519 PRELIMINARY; PRT; 343 AA.										
AC	092519,									
DT	01-FEB-1997 (TREMBLrel. 02, Created)									
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)									
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)									
DE	G83955 (G83955 protein).									
OC	Homo sapiens (Human).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;									
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
OX	NCBI_Taxid=9606;									
RP	[1]									
RC	SEQUENCE FROM N.A.									
RC	TISSUE=Cancellous bone;									
RA	Ohno I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.,									
RL	Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	TISSUE=uterus;									
RA	Strubeberg R.;									
DR	Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.									
DR	EMBL; D87119; BAA1350.1; -.									
DR	EMBL; BC002637; AAO02637.1; -.									
DR	HSSP; G63450; IA06.									
DR	InterPro; IPR00719; Prot_kinase.									
DR	Pfam; PF00069; Pfkinase; 1.									
DR	ProDom; PD000001; Prot_kinase; 1.									
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.									
DR	ATP-binding; Transferrase.									
SQ	SEQUENCE 343 AA; 38800 MW; BF8B7366DACB84FA CRC64;									
 Query Match 39.8%; Score 753.5; DB 4; Length 343; Best Local Similarity 47.3%; Pred. No. 2.9e-55; Matches 167; Conservative 50; Mismatches 115; Indels 21; Gaps 6										
Oy	2	RATPLAARAGSLSSKKRLLEDNDLDTBRPVQKARSGOPRLPCLLPSPPTADRA	TA 61							
Dd	5	RSTPTIARVGRSHNRKTDFEE--LSIRSAB-----PSQSFPMLGSPSPETPNLSHC	57							
Oy	62	VATASRLGPYLLAPBEGGARAYRALHCTGTGYTCCKVPV---QBALVLBPYALPPHK	118							
Dd	58	V---SCIGKYLLPLEBGDNHFRVAVHLHSGBELVCVKPFDISCYESLA---PCFLSAS	111							
Oy	119	HVARPTVIAGTOLLVAFFTTTHDDMSLVSRHRIPERBAVLPRONATALAHCHQGL	178							
Dd	112	NINOTTEILIGETAAYYPFERSYGDMHSFATCCKLRBBEARLYOJASNAHCHDGL	171							
Oy	179	VLRLDKLCRFVAFADRRKRKLVLNMLEDSCVLTGPDSDLMDXACPAVYGGPIILSRASYS	238							
Dd	172	VLRLDKRKRFKFOBERRTVLESLEDAYLIRGDODDSLXKHGCCPAYSPRILTNSGSYS	231							
Oy	239	GKAADWSLGVALTMTLAGHYPPQDSEVULLFGKTRRCAVALPAGLSAPACLVRCLR	298							
Dd	232	GKAADWSLGVALTMTLVGRYPFDHIEPSSLFSKIRQCQFNIPELTSKACCLRSILRR	291							
Oy	299	EPARBLTAGTILLHPWLRODPMPLAPTSSHLMBAOVVPGDGLDEAREEBCD	351							
Dd	292	EPERLISQBIILDPHFSTD-----FSVSNSGCGAKESVDQLVPDYMNENLD	339							
 RESULT 3 092519 PRELIMINARY; PRT; 343 AA.										
AC	092519,									
DT	01-FEB-1997 (TREMBLrel. 02, Created)									
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)									
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)									
DE	G83955 (G83955 protein).									
OC	Homo sapiens (Human).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;									
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
OX	NCBI_Taxid=9606;									
RP	[1]									
RC	SEQUENCE FROM N.A.									
RC	TISSUE=Cancellous bone;									
RA	Ohno I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.,									
RL	Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	TISSUE=uterus;									
RA	Strubeberg R.;									
DR	Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.									
DR	EMBL; D87119; BAA1350.1; -.									
DR	EMBL; BC002637; AAO02637.1; -.									
DR	HSSP; G63450; IA06.									
DR	InterPro; IPR00719; Prot_kinase.									
DR	Pfam; PF00069; Pfkinase; 1.									
DR	ProDom; PD000001; Prot_kinase; 1.									
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.									
DR	ATP-binding; Transferrase.									
SQ	SEQUENCE 343 AA; 38800 MW; BF8B7366DACB84FA CRC64;									
 Query Match 39.8%; Score 753.5; DB 4; Length 343; Best Local Similarity 47.3%; Pred. No. 2.9e-55; Matches 167; Conservative 50; Mismatches 115; Indels 21; Gaps 6										
Oy	2	RATPLAARAGSLSSKKRLLEDNDLDTBRPVQKARSGOPRLPCLLPSPPTADRA	TA 61							
Dd	5	RSTPTIARVGRSHNRKTDFEE--LSIRSAB-----PSQSFPMLGSPSPETPNLSHC	57							
Oy	62	V								

```

Db      112 NINQITEILGETKAYVFERSYGDMHSFVRTCKCLREBEARLFTQIAGAVAHCHDGL 171
Qy      179 VLRLDLKCRFPVADRERKULVLENLPSCVLTGPDDSLMKXKCAPVAGREILSSRASY 238
      172 VLRLDLKCRFPVADRERKULVLENLPSCVLTGPDDSLMKXKCAPVAGREILSSRASY 231
Qy      239 GKADVMSLQVLAFTMLAGHYPPQDSEPVLLFGKIRGAVLAPAGLSAPARCLVRCILRR 298
      232 GKADVMSLQVLAFTMLAGHYPPQDSEPVLLFGKIRGAVLAPAGLSAPARCLVRCILRR 291
Qy      299 EPAERLTATGILLHPMLRQDPMPPLAPTRSHLWBAQVVPDGLGDEAREEGD 351
      292 EPERLTSQELLDHPWFSTD-----FSVNSNGFKAKEVSDQLVPMMEENLD 339

RESULT 4
Q8K4K3 PRELIMINARY; PRT; 343 AA.
ID      08K4K3
AC      08K4K3
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10090;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      K18a-Tsch B., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA      Bagstaff S.M., Wyllie D.H., Harte M., O'Neill L.A.J., Quarnstrom E.B.,
RA      Dower S.K.;
RT      "Mammalian homologs of Drosophila tribbles (trib) control mitogen
RT      activated protein kinase signaling."
RL      Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF358867; AA45477.1; -.
DR      InterPro: IPR000719; Prot kinase.
DR      InterPro: IPR002290; Ser_Thr_kinase.
DR      InterPro: IPR001245; Tyr_kinase.
DR      Pfam; PF00069; Pkinase; 2.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      SMART; SM00220; S_TKc; 1.
DR      SMART; SM00219; TyKc; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM      ATP-binding; Transferase.
SQ      SEQUENCE 343 AA; 38758 MW; 0B3965B8B2087D74 CRC64;

Query Match      39.8%; Score 753; DB 11; Length 343;
Best Local Similarity 46.6%; Pred. No. 3.1e-55;
Matches 163; Conservative 49; Mismatches 108; Indels 30; Gaps 6;

Qy      2 RATPLAAPAGSLRKKRLLELDNLDTERPVQKARSGPQRLPCLPLSPPTAPRATA 61
      5 RSTPLIARVGRNRNTQDFE-----LSIRSAE-----PSQSFSPNIGSPEPPTPIVLSHC 57
Db      172 VLRLDLKCRFPVADRERKULVLENLPSCVLTGPDDSLMKXKCAPVAGREILSSRASY 238
Qy      62 VVATSRIGPVVLEPEEGGAVYRALHCPGTGYTKYVY---QBALVLEPVYARLPKH 118
      58 V---SGIGKYLLEPEEGGAVYRALHCPGTGYTKYVY---QBALVLEPVYARLPKH 111
Qy      119 HVAAPTEVLATQQLVLAFTTRTHGDMHSLVRSRRRIPEBAVLFROMATALAHCHQHG 178
      112 NINQITEILGETKAYVFERSYGDMHSFVRTCKCLREBEARLFTQIAGAVAHCHDGL 171
Db      179 VLRLDLKCRFPVADRERKULVLENLPSCVLTGPDDSLMKXKCAPVAGREILSSRASY 238
      172 VLRLDLKCRFPVADRERKULVLENLPSCVLTGPDDSLMKXKCAPVAGREILSSRASY 231
Qy      239 GKADVMSLQVLAFTMLAGHYPPQDSEPVLLFGKIRGAVLAPAGLSAPARCLVRCILRR 298
      232 GKADVMSLQVLAFTMLAGHYPPQDSEPVLLFGKIRGAVLAPAGLSAPARCLVRCILRR 291
Qy      299 EPAERLTATGILLHPMLRQDPMPPLAPTRSHLWBAQVVPDGLGDEAREEGD 351
      292 EPERLTSQELLDHPWFSTD-----FSVNSNGFKAKEVSDQLVPMMEENLD 339

```

```

Db      292 EPERLTSQELLDHPWFSTD-----FSVNSNGFKAKEVSDQLVPMMEENLD 339

RESULT 5
Q9EOL6 PRELIMINARY; PRT; 364 AA.
ID      Q9EOL6
AC      Q9EOL6
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      G-protein-coupled receptor induced protein G1G2 (fragment).
GN      G1G2.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxId=10116;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      STRAIN-Sprague-Dawley; TISSUE=Brain;
RA      Mayhaus M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,
RA      Nitsch R.M.;
RT      "Identification of a novel nuclear factor G1G2, as an m1-acetylcholine
RT      receptor-induced gene."
RL      Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF205438; AA35664.1; -.
DR      InterPro: IPR000719; Prot_kinase.
DR      Pfam; PF00069; Pkinase; 1.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM      ATP-binding; Receptor; Transferase.
FT      NON TER 1
SQ      SEQUENCE 364 AA; 40377 MW; 8F9167FB76DFCD37 CRC64;

Query Match      39.6%; Score 749.5; DB 11; Length 364;
Best Local Similarity 46.2%; Pred. No. 6.6e-55;
Matches 163; Conservative 46; Mismatches 119; Indels 25; Gaps 5;

Qy      2 RATPLAAPAGSLRKKRLLELDNLDTERPVQKARSGPQRLPCLPLSPPTAPRATA 61
      10 RGPALLFPARGTPARL-----LDVDDAAVAAKC---PRLSNCNPPYLLSPGSPCS 61
Db      106 AVLEPYARLPKHGVARPTEVLATQQLVLAFTTRTHGDMHSLVRSRRRIPEBAVLPRQ 165
      122 DKIRPYTQLPSHRNITGIVEVILGSKAYVFEKDFGDMHSYVSRKRLREBEARLFTQ 181
Qy      166 MATALAHCHQHGVLVRLDLKCRFPVADRERKULVLENLPSCVLTGPDDSLMKXKCAP 225
      182 IYSAVAHCHQSAVLVDLKLRFVFTSEBRTQRLSLBDTHMIKGBDLSXHGCPAY 241
Qy      226 VGEPIILSRASYSGKADVMSLQVLAFTMLAGHYPPQDSEPVLLFGKIRGAVLAPAGLS 285
      242 VGEPIILSRASYSGKADVMSLQVLAFTMLAGHYPPQDSEPVLLFGKIRGAVLAPAGLS 285
Db      286 APARCLVRCILRRBEARLTATGILLHPMLRQDPMPPLAPTRSHLWBAQVVPD 338
      302 PARCLVRCILRRBEARLTATGILLHPMLRQDPMPPLAPTRSHLWBAQVVPD 353
Qy      302 PARCLVRCILRRBEARLTATGILLHPMLRQDPMPPLAPTRSHLWBAQVVPD 338
      302 PARCLVRCILRRBEARLTATGILLHPMLRQDPMPPLAPTRSHLWBAQVVPD 353

RESULT 6
Q96RUB PRELIMINARY; PRT; 372 AA.
ID      Q96RUB
AC      Q96RUB
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      SKIPL.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

```

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Klee-Toth B., Wylie D.H., Ovarstrom E.E., Dower S.K.;
 RT "Identification of pro-inflammatory cytokine signaling network
 components by transcription expression screening.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF250310; AAK58174.1;
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; PKinase; 1.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 372 AA; 40980 MW; 3E2B5C87A4F98FDB CRC64;

Query Match 39.3%; Score 742.5; DB 4; Length 372;
 Best Local Similarity 46.8%; Pred. No. 1.8e-54;
 Matches 162; Conservative 47; Mismatches 126; Indels 11; Gaps 6;

OY 2 RALPLAAPA-GSLSRKRLDLD--NIDTERPVQKARSQP---QPRLPCLLPSPPTA 55
 DB 17 RGPALLFPATRGVPAKRLDADDAVAACPRLSGSSPPDYLSFGSPFC-SFQPPAA 75
 OY 56 PDATAVATA---SRGPVYLBPBEGGAYRALHCPGTGYTCQVYPVQALAVLEPYA 112
 DB 76 PGAGGSGSAPGSPRIADYLLPLAREBHVSRLCHTGRLECKVPIKHYQDKIRPYI 135
 OY 113 RLPPHKVAPRTVLAQTOLLYAFPTTHGDMSLVSRNRIPPEPAVLPRQATLAAH 172
 DB 136 QLPESHNTIGVIVETAYVFFPKSFGDMHSYVSRKRLREBAARLFKOIVSAVAH 195
 OY 173 CHQGLVLRDLKCRFPFADBERKKLVLENLSDCVLTGPDSDMDKACPAVYGPPILS 232
 DB 196 CHGSAIVLGDILKARKFPSTBERTQLRLSLDTHIKGDDALSDHGCPAYVSPILN 255
 OY 233 SRASYSKADAVMSLGVALLFTMLAGHYPFODSEPVLLFGKIRGAYVALPAGLSAPARCLV 292
 DB 256 TTGTYSKADAVMSLGVALLFTMLAGHYPFODSEPVLLFGKIRGAYVALPAGLSAPARCLV 315
 OY 293 RCLLRBPAPRLTATGILLHPMLRQDMPPLAPTRSHLMEAAQVVPD 338
 DB 316 RSLLRBPAPRLTATGILLHPMLRQDMPPLAPTRSHLMEAAQVVPD 360

RESULT 7

O9H2Y8 PRELIMINARY; PRT; 372 AA.

AC O9H2Y8;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DR G-protein-coupled receptor induced protein G1G2.
 GN G1G2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Mayhew M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,
 RA Nitsch R.M.;
 RT "Identification of a novel nuclear factor G1G2, an ml-acetylcholine
 receptor-induced gene.";
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF205437; AAG35663.1;
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; PKinase; 1.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Receptor; Transferase.
 SQ SEQUENCE 372 AA; 41008 MW; 5F54B50924B1365B CRC64;

Query Match 39.3%; Score 742.5; DB 4; Length 372;
 Best Local Similarity 46.8%; Pred. No. 2.6e-54;
 Matches 162; Conservative 46; Mismatches 127; Indels 11; Gaps 6;

OY 2 RALPLAAPA-GSLSRKRLDLD--NIDTERPVQKARSQP---QPRLPCLLPSPPTA 55
 DB 17 RGPALLFPATRGVPAKRLDADDAVAACPRLSGSSPPDYLSFGSPFC-SFQPPAA 75
 OY 56 PDATAVATA---SRGPVYLBPBEGGAYRALHCPGTGYTCQVYPVQALAVLEPYA 112
 DB 76 PGAGGSGSAPGSPRIADYLLPLAREBHVSRLCHTGRLECKVPIKHYQDKIRPYI 135
 OY 113 RLPPHKVAPRTVLAQTOLLYAFPTTHGDMSLVSRNRIPPEPAVLPRQATLAAH 172
 DB 136 QLPESHNTIGVIVETAYVFFPKSFGDMHSYVSRKRLREBAARLFKOIVSAVAH 195
 OY 173 CHQGLVLRDLKCRFPFADBERKKLVLENLSDCVLTGPDSDMDKACPAVYGPPILS 232
 DB 196 CHGSAIVLGDILKARKFPSTBERTQLRLSLDTHIKGDDALSDHGCPAYVSPILN 255
 OY 233 SRASYSKADAVMSLGVALLFTMLAGHYPFODSEPVLLFGKIRGAYVALPAGLSAPARCLV 292
 DB 256 TTGTYSKADAVMSLGVALLFTMLAGHYPFODSEPVLLFGKIRGAYVALPAGLSAPARCLV 315
 OY 293 RCLLRBPAPRLTATGILLHPMLRQDMPPLAPTRSHLMEAAQVVPD 338
 DB 316 RSLLRBPAPRLTATGILLHPMLRQDMPPLAPTRSHLMEAAQVVPD 360

RESULT 8

O91W04 PRELIMINARY; PRT; 372 AA.

AC O91W04;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DR Similar to phosphoprotein regulated by mitogenic pathways.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The PANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; BC006800; AA06800.1;
 DR EMBL; AK028626; BAC26038.1;
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; PKinase; 1.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 372 AA; 41281 MW; AD29BBA640B4862 CRC64;

Query Match 39.3%; Score 742.5; DB 11; Length 372;
 Best Local Similarity 46.7%; Pred. No. 2.6e-54;
 Matches 162; Conservative 44; Mismatches 119; Indels 25; Gaps 6;

OY 2 RALPLAAPAGSLSRKRLDLDNLTBRPVQKARSQPQPRLPCLIP---ISPPTAP-- 56
 DB 17 RGPGLLPAPARCTPARKL-----LDTD---DAGAVAAKCPRLSGSSPPDYLSFGSPFC 68
 OY 57 -----DRAVAVATARLGP-----YVLBPBEGGAYRALHCPGTGYTCQVYPVQAL 105

```

Db      69 POPPSTGTGSGSVSSGSPSRADYLLPLREHVSRLCIHTGRELRCRFPKIQY 128
      106 AVLEPYALRPKHVARPTVLAGTQLYAPFTRHGMHSLVSRHRIPEBAVLRQ 165
      129 DKIRPYQLPFSHNTGIVEVLLGSKAYVPEKPGFMSYVSRKRLREBAARLFKQ 188
      166 MATAALAHCHQGLVLRDLKCRFVPADREBKVLVLENTEDSCVLGPDDSLMDKACPAY 225
      189 IVSAVAHCHQSAIVLGDILKRFVSTERQTLRESIEDTHIIGEDDALSDDKGCAY 248
      226 VGPETLSSRASYSCKADVMSLGVLFMLAGHYRPODEPYLLFGKIRGAYALPAGLS 285
      249 VSPETLNTGTYSCKADVMSLGVLYTLVGRYPFHDSDPALPSKIRGQFCIPHYVS 308
      286 APACLVRCILRRPBARLTATGILLHPMLRQDMPPLAPTSHTLMEAAQVVPD 338
      309 PKARCLIRSLRRPSEBRLTAPQILLHPFVYLPR-GYVDSIGTSDQIYVE 360

```

RESULT 9

```

ID      08K4K4      PRELIMINARY;      PRT;      372 AA.
AC      08K4K4;
DT      01-OCT-2002 (TREMBLrel. 22, Created)
DT      01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      TRB-1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Klab-Tsch E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA      Baggett S.M., Wylie D.H., Harte M., O'Neill L.A.J., Owarnestrom B.E.,
RA      Dower S.K.;
RT      "Mammalian homologs of Drosophila cribbles (htrb) control mitogen
RT      activated protein kinase signaling."
RL      Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL, AF358866; AAA45478.1;
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_Chr_kinase.
DR      Pfam; PF00069; Pkinase; 2.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      SMART; SM00220; S_TKc; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM      ATP-binding; transferase.
SQ      SEQUENCE 372 AA; 41282 MW; 3A3DB82B46CD907F CRC64;

```

Query Match 38.6%; Score 729.5; DB 11; Length 372;
 Best Local Similarity 46.2%; Pred. No. 3.3e-53;
 Matches 163; Conservative 44; Mismatches 121; Indels 25; Gaps 6;

```

      2 RATPRLAPAGSLSRKKLELDNDLTERPVQKARSQPOPRLPCLLP--LSPTAP-- 56
      17 RGPCLLPFAKRGTAQRL-----LDTD---DAGVAAKCPRLSSCPDVLSPGSPCS 68
      57 -----DATAVAATASRLP-----YVLEBEGGKAYRALHCTGTGYTCVTPVQDAL 105
      69 POPPSTGTGSGSVSSGSPSRADYLLPLREHVSRLCIHTGRELRCRFPKIQYQ 128
      106 AVLEPYALRPKHVARPTVLAGTQLYAFTRHGMHSLVSRHRIPEBAVLRQ 165
      129 DKIRPYQLPFSHNTGIVEVLLGSKAYVPEKPGFMSYVSRKRLREBAARLFKQ 188
      166 MATAALAHCHQGLVLRDLKCRFVPADREBKVLVLENTEDSCVLGPDDSLMDKACPAY 225
      189 IVSAVAHCHQSAIVLGDILKRFVSTERQTLRESIEDTHIIGEDDALSDDKGCAY 248
      226 VGPETLSSRASYSCKADVMSLGVLFMLAGHYRPODEPYLLFGKIRGAYALPAGLS 285
      249 VSPETLNTGTYSCKADVMSLGVLYTLVGRYPFHDSDPALPSKIRGQFCIPHYVS 308

```

```

      286 APACLVRCILRRPBARLTATGILLHPMLRQDMPPLAPTSHTLMEAAQVVPD 338
      309 PKARCLIRSLRRPSEBRLTAPQILLHPFVYLPR-GYVDSIGTSDQIYVE 360

```

RESULT 10

```

ID      08R2V8      PRELIMINARY;      PRT;      257 AA.
AC      08R2V8;
DT      01-JUN-2002 (TREMBLrel. 21, Created)
DT      01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Hypothetical 29.2 kDa protein (Fragment).
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Strauberg R.;
RL      Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC027159; AAH27159.1;
DR      InterPro; IPR000719; Prot_kinase.
DR      Pfam; PF00069; Pkinase; 1.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM      Hypothetical protein; ATP-binding; transferase.
FT      NON TRR 1
SQ      SEQUENCE 257 AA; 29232 MW; 84BAC0DD4767F51 CRC64;

```

Query Match 34.5%; Score 653; DB 11; Length 257;
 Best Local Similarity 52.3%; Pred. No. 5.7e-47;
 Matches 134; Conservative 34; Mismatches 68; Indels 20; Gaps 3;

```

      96 CKYVPV---QBALVLEPYALRPKHVARPTVLAGTQLYAPFTRHGMHSLVSRH 152
      3 CKVEISICYGSLA---PCFLSHNSINQTTETILSTKAYVFERSYGDMHSEYRICK 59
      153 RIPEBAVLRPQATYALAHCHQGLVLRDLKCRFVPADREBKVLVLENTEDSCVLTP 212
      60 KLRBEBAARLFYQIASVAHCHQGLVLRDLKRFKDEBRTRVKLSLEDAVILRGD 119
      213 DDSLMDHGCAPYVGPETLSSRASYSCKADVMSLGVLFMLAGHYRPOSEPYLLFGK 272
      120 DDSLSDHGCAPYVSPETLNTGSGYSCKADVMSLGVLYTLVGRYPFHDIPSSLSK 179
      273 IIRGAYALPAGLSAPACLVRCILRRPBARLTATGILLHPMLRQDMPPLAPTSHTLME 332
      180 IIRGQPMIFETLSKAKCLIRSIIRKREPERLTSQELIDHPFSTD----- 225
      333 AQVVDGLGLDEAREE 348
      226 FSVNSGFGAKEACDQ 241

```

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=98000262; PubMed=9342215;
 RA Wilkin P., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,
 RA Dumont J.B., Maenhaut C.;
 RT "Characterization of a phosphoprotein whose mRNA is regulated by the
 RT mitogenic pathways in dog thyroid cells";
 RL Eur. J. Biochem. 248:660-669(1997).
 DR EMBL; AJ000480; CAA04119.1; -
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KM ATP-binding; Transferase.
 FT NON TER 1
 SQ SEQUENCE 224 AA; 25479 MW; 740CD0905F86499B CRC64;
 Query Match 32.3%; Score 611.5; DB 4; Length 224;
 Best Local Similarity 56.7%; Pred. No. 1.5e-43;
 Matches 119; Conservative 31; Mismatches 59; Indels 1; Gaps 1;
 QY 129 GTQLLYAFPTTRGMSLVSRSHRIPEPEAAVLFROMATALAHCHQGLVLDLTKRF 188
 DB 4 GETKAYFFPKSGDMHSTVRSRKLREBAARLKQIVSAVAHCHQSAIVLGLTKRF 63
 QY 189 VPADREKKLVLENDSCVLTPDDSLMDKHA CPAVYGPRISSRASVGRADVWSIG 248
 DB 64 VPTSERQRLRESLEDTHIMKGBDALSCKGCPAVYSPILNTGTGSGRADVWSIG 123
 QY 249 VALPTMLAGHPQSESPVLLFGKTRRGAVALPAGLSAPARCLVRLRREAEKLTATG 308
 DB 124 VMLVTLVGRYFPHSDSPALPSKIRRGFCRPIHSIPARCLIRSLRREPSERTLAE 183
 QY 309 ILHPLMGDDPWLAPTRSHLWEAAOVVPD 338
 DB 184 ILHWPESVLEP-GYIDSEIGTSQIVPE 212
 RESULT 12
 QV321 PRELIMINARY; PRT; 484 AA.
 AC QV321;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Putative TRBL protein (CG5408 protein).
 GN TRBL OR CG5408.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.B., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baker B.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benson P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brodeur P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadietu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo S., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,

RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Houtlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mikhina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier B., Spreading A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weinstein J.,
 RA Williams S.M., Woodge T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye U., Yeh R.-P., Zaveri U.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN
 RP SEQUENCE FROM N.A.
 RA Seher T.C., Leptin M.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB003591; AAF51590.1; -
 DR EMBL; AF204688; AAF26374.1; -
 DR PDBase; PDB0028978; trbl.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KM Hypothetical protein; ATP-binding; Transferase.
 SQ SEQUENCE 484 AA; 54077 MW; 3E3BD3E564580D7 CRC64;
 Query Match 23.0%; Score 435; DB 5; Length 484;
 Best Local Similarity 33.4%; Pred. No. 2.5e-28;
 Matches 116; Conservative 45; Mismatches 134; Indels 52; Gaps 8;
 QY 31 VOKRSGOPRLPPCLPLSPPTARDATNAVATSRGPLYLREBSGRVRLHCT 90
 DB 104 IQQRILISAQPSHISAATAATP-ASYRHLVDLTASN- - - - -RCVDLFT 147
 QY 91 GTEYTCYVYQVAL-AVLEPYARLPKHVARPEVLA- - - - -GTOL 133
 DB 148 GEQFLCR--VNEPLHKQRAYFQIQCHDELRSSTIYGHPLRPVNDIIPLTQRTYIL 205
 QY 134 YA-----PFTTHGDMHSLVSRSHRIPEPEAAVLFROMATALAHCHQGLVLDL 183
 DB 206 IAPVQERDSTGVTGVVENHTIYIRHAKRLCETEARAI FQIQCTGVQVCHRNGLIIDL 265
 QY 184 KLCPVPADRRKKLVLENDSCVLTPDDSLMDKHA CPAVYGPRISSRASVGRADV 243
 DB 266 KLRKFFYIDARKTLQVHESLGSMTLDSEDTLSDKICPPLTVABELCPQDTYKGPAD 325
 QY 244 VMSGLVALFTMLAGHPQSESPVLLFGKTRRGAVALPAGLSAPARCLVRLRREAEK 303
 DB 326 MMSGLVILYTLVQGYPRYERANCNLITIRHGNVQIPLTLSKSRWLLSLRLDYER 385
 QY 304 LTANGILHPLVLRQDDPWLAPTRSHLWEAAOVVPDGLDARABEG 350
 DB 386 MTASHILFPLWLRQ---RPFHMYLPVDVVAED---WSDARDEG 425
 RESULT 13
 QV327 PRELIMINARY; PRT; 218 AA.
 AC QV327;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Similar to phosphoprotein regulated by mitogenic PATHWAYS

DE homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RA [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;

RX MEDLINE=22354683; PubMed=12466851;

RA The RIKEN Consortium;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs";

RL Nature 420:563-573(2002).

DR EMBL; AK040738; BAC30688.1; -

DR EMBL; AK041212; BAC30685.1; -

SQ SEQUENCE 218 AA; 23907 MW; P3A5AFB16E3450A8 CRC64;

Query Match 17.2%; Score 326; DB 11; Length 218;

Best Local Similarity 40.0%; Pred. No. 1.4e-19;

Matches 84; Conservative 23; Mismatches 79; Indels 24; Gaps 5;

QY 2 RATPLAAGSLSRKRLLELDNLDTERPVQKARSGPQRLPPCLLP---LSPTAP-- 56

DB 17 RGPGLFPARGPAPKRL-----LDTD---DAGVAKACRLSCLSSPPDVLSPGSPCS 68

QY 57 -----DRATVATASRLGP-----YLLPEBEGRAYRALHCTGTCTKYVQBAL 105

DB 69 POPPSTQGTGSCVSPGSPSRADYLLPLAERHVSRLCITGTRELRCKEPPIKYQ 128

QY 106 AVLEPYARLPKHGVAPTEVLACTQLLYAFPTTHGDMSLVSRRIEPPBAVLPQ 165

DB 129 DKIRPYQLPSHSNITGIYVLLGSKAYVFFEKDFODMSYVSRRLRBEBAARLFQ 188

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 189 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 218

Search completed: January 15, 2004, 14:58:44

Job time: 43 secs

QY 57 -----DRATVATASRLGP-----YLLPEBEGRAYRALHCTGTCTKYVQBAL 105

DB 230 POPPSTQGTGSCVSPGSPSRADYLLPLAERHVSRLCITGTRELRCKEPPIKYQ 289

QY 106 AVLEPYARLPKHGVAPTEVLACTQLLYAFPTTHGDMSLVSRRIEPPBAVLPQ 165

DB 290 DKIRPYQLPSHSNITGIYVLLGSKAYVFFEKDFODMSYVSRRLRBEBAARLFQ 349

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

QY 166 MATAAHCHQHGLVLRDLKLCRFVADRRER 195

DB 350 IVSAVAHCHQSAIVLGDLLKRLKRVFSTERR 379

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:55:46 ; Search time 18 Seconds

(without alignments)
935.309 Million cell updates/sec

Title: US-09-909-474D-2

Perfect score: 1891
Sequence: 1 MKATPLAAPAGSLSRKKRLK.....GLGLDPAKRESEGDREVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1887	99.8	358	1 NIPK_HUMAN	Q96R17 homo sapien
2	1345	71.1	354	1 NIPK_MOUSE	Q9K4K2 mus musculu
3	1331.5	70.4	349	1 NIPK_RAT	Q9W2G6 rattus norv
4	309	16.3	619	1 SNPI_CANTR	Q94168 candida tro
5	304.5	16.1	646	1 CNK_HUMAN	Q944B4 homo sapien
6	304	16.1	620	1 SNPI_CANAL	P52497 candida alb
7	298.5	15.8	631	1 CNK_MOUSE	Q60806 mus musculu
8	291	15.4	615	1 CNK_MOUSE	Q91011 rattus norv
9	287	15.2	1518	1 KKKI_YEAST	P54244 saccharomyc
10	284.5	15.0	611	1 KKKI_YEAST	Q00372 candida gla
11	281	14.9	512	1 KIILO_ARATH	Q38997 arabidopsis
12	274.5	14.5	633	1 SNPI_YEAST	P06782 saccharomyc
13	273.5	14.5	533	1 KCCD_RAT	P57991 rattus norv
14	271.5	14.4	593	1 CDR1_SCHPO	P07334 echinosacch
15	271	14.3	576	1 SNPI_SCHPO	Q74536 echinosacch
16	269	14.2	542	1 KCCB_MOUSE	P28652 mus musculu
17	268	14.2	632	1 YNA3_CABEL	P45894 caenorhabdi
18	266.5	14.1	499	1 KCCD_HUMAN	Q13357 homo sapien
19	266	14.1	542	1 KCCB_HUMAN	P08413 rattus norv
20	266	14.1	664	1 KCCB_HUMAN	Q13554 homo sapien
21	263	13.9	1142	1 GIN4_YEAST	Q12263 saccharomyc
22	261.5	13.8	552	1 AAK2_HUMAN	P54646 homo sapien
23	261.5	13.8	891	1 KINI_SCHPO	P22987 echinosacch
24	260.5	13.8	552	1 AAK2_RAT	Q92017 rattus norv
25	260.5	13.8	766	1 SNIL_HUMAN	P57059 homo sapien
26	258	13.6	502	1 RKI1_SECCE	Q02773 secale cere
27	256	13.5	610	1 CDP1_ARATH	Q06850 arabidopsis
28	255.5	13.5	478	1 KCCA_HUMAN	Q9UQ77 homo sapien
29	255.5	13.5	478	1 KCCA_HUMAN	P11735 rattus norv
30	252.5	13.4	433	1 STKB_HUMAN	Q15831 homo sapien
31	250	13.2	735	1 K6A1_RAT	Q9NYJ3 rattus norv
32	247.5	13.1	685	1 SNK_HUMAN	Q9NYJ3 homo sapien
33	247	13.1	332	1 AAIIP_WHEAT	Q02066 triticum ae

34	246.5	13.0	724	1 K6A1_MOUSE	P18653 mus musculu
35	244.5	12.9	478	1 KCCA_MOUSE	P11798 mus musculu
36	244.5	12.9	735	1 K6A1_HUMAN	Q15418 homo sapien
37	243.5	12.9	472	1 KCCG_HUMAN	Q13555 homo sapien
38	243.5	12.9	527	1 KCCG_RAT	P11730 rattus norv
39	243.5	12.9	529	1 KCCG_MOUSE	Q93319 mus musculu
40	242.5	12.8	353	1 ASK2_ARATH	P43292 arabidopsis
41	242	12.8	661	1 Y537_HUMAN	O60285 homo sapien
42	242	12.8	915	1 KCC4_YEAST	P25389 saccharomyc
43	241	12.7	542	1 CDP3_ORYSA	P53684 oryza sativ
44	240.5	12.7	548	1 AAK1_RAT	P54645 rattus norv
45	240	12.7	682	1 SNK_MOUSE	P53351 mus musculu

ALIGNMENTS

RESULT 1
ID NIPK_HUMAN STANDARD; PRT; 358 AA.
AC Q96R17; Q9TA16; Q9H5M8; Q9NUD2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DS Neuronal cell death inducible putative kinase (SKRP3).
GN NIPK OR C20ORP97.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
[1]
RN RP SEQUENCE FROM N.A.
RA Kish-Toth E., Wylie D.H., Qvarnstrom E.E., Dower S.K.;
RT "Identification of pro-inflammatory cytokine signaling network
components by transcription expression screening."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN RP SEQUENCE FROM N.A.
RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Oktani T., Ota T., Suzuki Y., Oiyashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegaki T., Sugano S.;
RT "NBD human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RX MBLIRB-21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stevridou G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blake S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Billington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle B., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lhvaesliho M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McElroy K., McMurtry A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prachinlang S.R., Plumb R.W., Ramsey H.,
RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Stuebe C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.B.,
RA Swann R.M., Syamone N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Whitling L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).

RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervix, and Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA DiCicco L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stetson M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Vadn T.B., Toshiyuki S., Carlini P., Prange C.,
 RA Bata S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.M.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J., Schmitz J., Myers R.M.,
 RA Rodriguez A.C., Grimwood J., Schmitz J., Skalska U., Smallus D.B.,
 RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: May play an important role in a common pathway leading
 CC to programmed neuronal cell death. Does not appear to function in
 CC the programmed death of non-neuronal cells. May serve as an
 CC endogenous antagonist competing for substrate with functional
 CC kinases that act to promote neuronal cell survival (by
 CC similarity).
 CC -1- SIMILARITY: Contains 1 protein kinase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation in
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-6db.ch/announce/>
 CC or send an email to license@1sb-6db.ch).
 CC -----
 DR EMBL: AF250311; AAK58175.1; -
 DR EMBL: AK026945; BAB15597.1; -
 DR EMBL: AL034548; CAB1634.1; -
 DR EMBL: BC019363; AAK19363.1; -
 DR EMBL: BC027484; AAK27484.1; -
 DR Genew; HGNC:16228; C200197;
 DR InterPro; IPR000719; Prot. kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot. kinase; 1.
 DR SMART; SM00220; 5_TKC;_KINASE_DOM; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Apoptosis.
 FT DOMAIN 68 316 PROTEIN KINASE.
 FT CONFLICT 84 84 Q -> R (IN REF. 4; AAK27484).
 FT CONFLICT 105 105 L -> P (IN REF. 2).
 FT CONFLICT 114 114 L -> V (IN REF. 1).
 FT CONFLICT 194 195 ER -> DREK (IN REF. 1).
 SO SEQUENCE 358 AA; 39577 MW; CE15F089A81BDD63 CAC64;
 Query Match 99.8%; Score 1887; DB 1; Length 358;
 Best Local Similarity 99.7%; Pred. No. 1e-140;
 Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRATPLAAGSLSKKRLLEDDNIDTERPVOKRARSPPQLPLPCLPLSPPTAPDRAT 60
 DB 1 MRATPLAAGSLSKKRLLEDDNIDTERPVOKRARSPPQLPLPCLPLSPPTAPDRAT 60
 QY 61 AVATASRLGPTVYLEPEREGRAVRLHGPCTGTYCKYQVQALAVLBPVAPRPHGV 120
 DB 61 AVATASRLGPTVYLEPEREGRAVRLHGPCTGTYCKYQVQALAVLBPVAPRPHGV 120
 QY 121 ARPTVLAGTQLYAFTRTHGDMHSLVRSRHRIRPEPAAVLFRQMATALAHCHQGLVL 180
 DB 121 ARPTVLAGTQLYAFTRTHGDMHSLVRSRHRIRPEPAAVLFRQMATALAHCHQGLVL 180

DB 121 ARPTVLAGTQLYAFTRTHGDMHSLVRSRHRIRPEPAAVLFRQMATALAHCHQGLVL 180
 QY 181 RDLFLCRFPVADRRERKKLVLENLSDCYLTGPDSDLMDKACPAVYGPBILSSPASYGK 240
 DB 181 RDLFLCRFPVADRRERKKLVLENLSDCYLTGPDSDLMDKACPAVYGPBILSSPASYGK 240
 QY 241 AADVMSLGVALFTMLAGHYPPQDSEPVLLPCKIRGVALPAGISAPRCLVRCILRRP 300
 DB 241 AADVMSLGVALFTMLAGHYPPQDSEPVLLPCKIRGVALPAGISAPRCLVRCILRRP 300
 QY 301 AERLTATGILHPLRQDPMPLAPTRSHLWAAQVPGDGLDREAREEGREVTLYG 358
 DB 301 AERLTATGILHPLRQDPMPLAPTRSHLWAAQVPGDGLDREAREEGREVTLYG 358
 RESULT 2
 NPK_MOUSE
 ID NPK_MOUSE STANDARD; PRT; 354 AA.
 AC Q8K4K2_Q92157; 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-FEB-2003 (Rel. 42, Last annotation update)
 DE Neuronal cell death inducible putative kinase (TRB-3).
 GN NPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Kles-Toth E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
 RA Bagstaff S.M., Wylie D.H., Harle W., O'Neill L.A.J., Owerheltrom E.B.,
 RA Dower S.K.;
 RT "Mammalian homologs of Drosophila trtb3s (trb) control mitogen
 RL activated protein kinase signaling."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kaeukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikielido I., Otsu N., Saito R., Suzuki H., Yamazaki I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Betalov S., Betsel K.M.,
 RA Blake J.A., Brad T., Brusec V., Chochia C., Corbani L.B., Cousins S.,
 RA Dalia E., Dragan T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gassterland T., Gariboldi M., Giesel C., Godzik A., Gough J.,
 RA Grimmond S., Guertlich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierzki R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Mik H.,
 RA Nagashima T., Namata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Yanagisawa M., Yang L.,
 RA Wilmfing L.G., Wyshaw-Boris A., Yanagisawa M., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hitozane-Kishikawa T., Kono H., Nakamura M., Sakazune N., Sato K.,
 RA Shiraki T., Waki K., Kawaji J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yaunushl A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shemen C.M., Schuler G.D.,

```

RA Altschul S. F., Zeeberg B., Buecaw K. H., Scheffer C. F., Bhat N. K.,
RA Hopfner R. F., Jordan H., Moore T. Max S. I., Wang J., Hsieh P.,
RA Datchenko L., Marinova K., Farmer A. A., Rubin G. M., Hong L.,
RA Stapleton M., Soares M. B., Bonaldo M. F., Casavant T. L., Schetz T. B.,
RA Brownstein M. J., Ueda T. B., Tohkiyuki S., Garinoti P., Prange C.,
RA Raha S. S., Loeuallan N. A., Peters G. J., Abramson R. D., Mullaly S. J.,
RA Boask S. A., McEwan P. J., McKernan K. J., Malek J. A., Gunaratne P. H.,
RA Richards S., Morley K. C., Hale S., Garcia A. M., Gay L. J., Hulyk S. W.,
RA Villalón D. K., Muzny D. M., Sodergren E. J., Lu X., Gibbs R. A.,
RA Pahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A. C., Shevchenko Y., Bouffard G. G.,
RA Blakesley R. W., Touchman J. W., Green E. D., Dickson M. C.,
RA Rodriguez A. C., Grimwood J., Schmutz J., Myers R. M.,
RA Buterfield Y. S. N., Krzywnski M. I., Skalska U., Smalls D. B.,
RA Scherch A., Schein J. E., Jones S. J. M., Marra M. A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -I- FUNCTION: May play an important role in a common pathway leading
CC to programmed neuronal cell death. Does not appear to function in
CC the programmed death of non-neuronal cells. May serve as an
CC endogenous antagonist competing for substrate with functional
CC kinases that act to promote neuronal cell survival (By
CC similarity).
CC -I- SIMILARITY: Contains 1 protein kinase domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC
DR EMBL, AF358868, AAM5476.1, -
DR EMBL, AK089931, BAC41002.1, -
DR EMBL, BC012955, AAM12955.1, -
DR InterPro: IPR000719, Prot. Kinase.
DR InterPro: IPR002290, Ser. Thr. Kinase.
DR Pfam: PF00069, Kinase, 2.
DR ProDom: PD000001, Prot. Kinase, 1.
DR SMART, SM00220, S_TKc, 1.
DR PROSITE, PS50011, PROTEIN_KINASE_DOM, 1.
KW Apoptosis.
FT DOMAIN 68 315 PROTEIN KINASE.
FT CONFLICT 157 157 S -> P (IN REF. 3).
FT CONFLICT 219 219 K -> T (IN REF. 1).
FT CONFLICT 239 264 MISSING (IN REF. 3).
FT CONFLICT 301 354 SERVALGILGLHPVIRBDHGSPSPQSDRRMDQVPPDQGL
FT LEAEAGEGVGLG -> CRATCGGNPLASLVGRSRLS
FT STV (IN REF. 2).
FT SEQUENCE 354 AA: 39022 MW: 2C8283FC119B59F CXC64;
50
Query Match 71.1%; Score 1345; DB 1; Length 354;
Best Local Similarity 74.0%; Pred. No. 3e-98;
Matches 265; Conservative 29; Mismatches 60; Indels 4; Gaps 2;
Oy 1 MRATPLAAPASGLSKKRLLEDDNIDTRTPRPOKARASGFORLPCLPLSPPTAPDRAT 60
Db 1 MRATPLAAPASGLSKKRLLEDDNIDTRTPRPOKARASGFORLPCLPLSPPTAPDRAT 60
Oy 61 AVATASRLAGPVVLEPEEGGAGAYRALHCPGTGEVTCCKVYPOEALAVLEPAARLPKHV 120
Db 61 AVATASRLAGPVVLEPEEGGAGAYRALHCPGTGEVTCCKVYPOEALAVLEPAARLPKHV 120
Oy 121 ARPEVLVAGTOLVAFPTRHGMHSLVRSRRIRPEPAAYLFRQMATALAHCHQGLVYL 180
Db 121 ARPEVLVAGTOLVAFPTRHGMHSLVRSRRIRPEPAAYLFRQMATALAHCHQGLVYL 180
Oy 181 RDLTKLRVPVADRRBKLVLENDSCVLTGPPDSIMDKACPAVYGPETILSPASVSGK 240
Db 181 RDLTKLRVPVADRRBKLVLENDSCVLTGPPDSIMDKACPAVYGPETILSPASVSGK 240

```

Qy	241	AADWVSLGVALFTWMAGHYPPDSDSPVLFGFIRRGAAVALPGLSGAPARCLVRCLLRSP	300
Dd	241	AADVWSLGVALLFTMLAGRYPPHDSBVLFFGKIRGTALPGSLAPARCLVRCLLRSP	300
Qy	301	ASRLVATGILLHPMLRQDPMDPLAPRSHLMENAAQVVPDGLGDREABRBGRDRVVLYG	358
Dd	301	SERVLVGLILHPMLREDHGVRSPQSDRRREMDQVVPDGPQLERA--EKG--EVGLYG	354
 RESULT 3 NIPK_RAT			
ID	NIPK_RAT	STANDARD;	PRT; 349 AA.
AC	O9WT06;		
Dt	28-FEB-2003	(Rel. 41, Created)	
Dt	28-FEB-2003	(Rel. 41, Last sequence update)	
Dt	28-FEB-2003	(Rel. 41, Last annotation update)	
Dt	Neuronal cell death inducible putative kinase.		
Gn	NIPK.		
Os	Rattus norvegicus (Rat).		
Oc	Euteleostomi; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Ox	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
Rn	NCBI_Taxid=10116;		
Rp	[1]		
Rc	SEQUENCE FROM N.A.		
Rx	TISUS=Neuron;		
Rx	MEDLINE=99262087; PubMed=10329375;		
Rt	Mayumi-Matsuda K., Kojima S., Suzuki H., Sakata T.;		
Rt	"Identification of a novel kinase-like gene induced during neuronal		
Rt	cell death."		
Rl	Biochem. Biophys. Res. Commun. 258:260-264(1999).		
Cc	-1- FUNCTION: May play an important role in a common pathway leading		
Cc	to programmed neuronal cell death. Does not appear to function		
Cc	in the programmed death of non-neuronal cells. May serve as an		
Cc	endogenous antagonist competing for substrate with functional		
Cc	kinases that act to promote neuronal cell survival.		
Cc	-1- TISSUE SPECIFICITY: Detected only in the lung. Not detected in the		
Cc	heart, brain, spleen, liver, skeletal muscle, kidney and testis.		
Cc	-1- INDUCTION: Expression induced during programmed cell death evoked		
Cc	in neuronal cells by NGF-depletion.		
Cc	-1- SIMILARITY: Contains 1 protein kinase domain.		
Cc	-----		
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration		
Cc	between the Swiss Institute of Bioinformatics and the EMBL outstation.		
Cc	The European Bioinformatics institute. There are no restrictions on its		
Cc	use by non-profit institutions as long as its content is in no way		
Cc	modified and this statement is not removed. Usage by and for commercial		
Cc	entities requires a license agreement (see http://www.isb-sdb.ch/announce		
Cc	or send an email to license@isb-sdb.ch).		
Cc	-----		
Dr	EMBL; AB020967; BAA7582.1; -		
Dr	InterPro; IPR000719; Prot_kinase.		
Df	Pfam; PF00069; Kinase; 2.		
Df	Prodom; PD000001; Prot_kinase; 1.		
Df	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.		
Kw	Apoptosis.		
Ft	DOMAIN		
Sq	SEQUENCE 349 AA; 38602 MM; 3050F9BFBE94CD815 CRC64;		
 PROTEIN KINASE.			
 Query Match 70.4%; Score 1331.5; DB 1; Length 349;			
 Best Local Similarly 73.5%; Pred. No. 3.4e-97;			
 Matches 263; Conservative 26; Mismatches 60; Indels 9; Gaps 3;			
Qy	1	MRATPLAAPGSLSRKRLSLEDDNLDERPRQKARSGGOPRLRPCLLPSPPARPDAT	60
Dd	1	MRATSLAASADVPCRKKPLRFDDNDIVDCFLKRVDRBPBGPTPSL-----PPASDISP	55
Qy	61	AVATASRLGTYVLLPEPEGRAAYRALCPGTBTCTCYVYPOBALAVLEPYARLPKHCV	120
Dd	56	AVAPATRLGYTLIERQGNCSTRALHCPTGTETTCNVYPASBAQDALAPAKLPTHOV	115
Qy	121	ARPLEVLAGQGLYAFTTRTHGMHSLVRSNRHIPSEBAVLFROMATALACHQHQGLVT	180
Dd	116	ARPLEVLAGSQGLLYFTFKTHGHLSVRSRGIPEBEAALFPQMSSAVVHCHRGILT	175

QY 181 RDLKLCFVFPADRRKKLVLENLDSCVLTGPDSDLMDKACPAVYGPETLSSPASYGK 240
 DB 176 RDLKLRFFVFNCRKTLVLENLDACTMTGPDSDLMDKACPAVYGPETLSSPASYGK 235
 QY 241 AADVMSIGVALFTMLAGHYPPQDSBPVLTGKIRGVALPAGLSAPARCLVRCILRRP 300
 DB 236 AADVMSIGVALFTMLAGHYPPQDSBPVLTGKIRGVALPAGLSAPARCLVRCILRRP 295
 QY 301 AERLTATGILHPMLRQDPMPALPTRSHLMEAAOVDPDGLDEARBEGDREVTLYG 358
 DB 296 SERVALGILHPMLRQDSCQVSPPRSDRREMDQVVPDGPQLBBA--EEG--EYGLYG 349

RESULT 4
 SNF1_CANTR STANDARD, PRT, 619 AA.
 AC 094168;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
 GN SNF1.
 OS Candida tropicalis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitospotic Saccharomycetales; Candida.
 OX NCBI_TaxID=5482;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanai T., Ogawa K., Ueda M., Tanaka A.;
 RT "Genetic evaluation of the function of SNF1 in *Candida tropicalis*,"
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
 CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
 CC PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB024535; BAA75889.1; -
 DR HSRP; Q63450; 1A06. Prot kinase.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KM Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 17 29 POLY-HIS.
 FT NP_BIND 52 303 PROTEIN_KINASE.
 FT NP_BIND 58 66 ATP (BY SIMILARITY).
 FT BINDING 81 81 ATP (BY SIMILARITY).
 FT ACT_SITE 174 174 BY SIMILARITY.
 FT MOD_RES 207 207 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SQ SEQUENCE 619 AA; 70323 MW; 0FCPIFC3DCET706D7 CRC64;

Query Match 16.3%; Score 309; DB 1; Length 619;
 Best Local Similarity 30.9%; Pred. No. 6.5e-17;
 Matches 95; Conservative 48; Mismatches 128; Indels 36; Gaps 12;
 QY 40 QPRLPCLPLSPPTAPDRATVATASRLGPLYLLEP-EEG--GRAVRLHCPGTGTEYC 96

DB 32 QPAP---IPTDVPNP-----ANRIGYQIIKTLGSSFGKRVKLAQVGTGQKVAL 80
 QY 97 KVVYQVALAVLE-----PYARLPFHGVAPRTEVLATGTLVAFPTRTGDMSL 147
 DB 81 KIIN-RKTLAKSDMQGVEREISYLRLLRPHHIIKLVDIVASKOIIIVISPAKELPDY 139
 QY 148 VRSHRRIPEPBAVALFPQMATALANQHOGCLVLDLKLCRVFPADREKKVLLENLEOSC 207
 DB 140 IVQRKMPEDBARFPQOIIAAVEYCHRHKIVHRDLKRENLLDDQLVKXADPGL--SN 197
 QY 208 VLTGPDSDLMDKAC--PAVYGPETLSSPASYGKAADVMSIGVALFTMLAGHYPPQDS 265
 DB 198 IWT---DGNFLKTSGSGSPNTAAPREVISKL-YAGPBDVMSGGVILYMLCGRLPFDDBF 253
 QY 266 PVLLFGKIRGVALPAGLSAPARCLVRCILRRPARELTATGILHPMLRQDPMP---L 322
 DB 254 IPLAFKKSINGVTLTPYLSGAGHLLTRMLVVPPLRITTHIMEDEWFPQD--WPDVLT 312
 QY 323 APTRSHL 329
 DB 313 PPDLSKI 319

RESULT 5
 CNK_HUMAN STANDARD, PRT, 646 AA.
 ID CNK_HUMAN
 AC Q9H4B4; Q15767;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-
 DE inducible kinase) (Proliferation-related kinase).
 GN CNK OR FNK OR PRK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TIGR=Embryo;
 RX MEDLINE=20493044; PubMed=11039900;
 RA Holtrich U., Wolf G., Yuan J., Bereliter-Hahn J., Karn T., Weiler M.,
 RA Kauegmann G., Renli M., Andreesen R., Kauegmann M., Kuhl D.,
 RA Streibhardt K.;
 RT "Adhesion induced expression of the serine/threonine kinase FNK in
 RT human macrophages.";
 RL Oncogene 19:4832-4839(2000).
 RN [2]
 RP SEQUENCE OF 28-646 FROM N.A.
 RC TIGR=Placenta;
 RX MEDLINE=96325053; PubMed=8702627;
 RA Li B., Ouyang B., Pan H., Reissmann P.T., Slamon D.J., Arceci R.,
 RA Lu L., Dai W.;
 RT "Prk, a cytokine-inducible human protein serine/threonine kinase whose
 RT expression appears to be down-regulated in lung carcinomas.";
 RL J. Biol. Chem. 271:19402-19408(1996).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=96019242; PubMed=9353331;
 RA Ouyang B., Pan H., Lu L., Li J., Stambrook P., Li B., Dai W.;
 RT "Human Prk is a conserved protein serine/threonine kinase involved in
 RT regulating M phase functions.";
 RL J. Biol. Chem. 272:28646-28651(1997).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=20027391; PubMed=10557092;
 RA Ouyang B., Li W., Pan H., Meadows J., Hoffmann I., Dai W.;
 RT "The physical association and phosphorylation of Cdc25C protein
 RT phosphatase by Prk.";
 RL Oncogene 18:6029-6036(1999).
 CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN REGULATING M
 CC PHASE FUNCTIONS DURING THE CELL CYCLE. MAY ALSO BE PART OF THE

CC SIGNALING NETWORK CONTROLLING CELLULAR ADHESION. IN VITRO, IS ABLE
 CC TO PHOSPHORYLATE CDC25C AND CASEIN.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBUNIT: BINDS TO THE CALCIUM/INTEGRIN-BINDING PROTEIN (CIB). THIS
 CC INTERACTION PROBABLY OCCURS VIA THE POLO-BOX DOMAIN.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.
 CC -1- TISSUE SPECIFICITY: TRANSCRIPTS ARE HIGHLY DETECTED IN PLACENTA,
 CC LUNG, FOLLOWED BY SKELETAL MUSCLE, HEART, PANCREAS, OVARIES AND
 CC KIDNEY AND WEAKLY DETECTED IN LIVER AND BRAIN. MAY HAVE A SHORT
 CC HALF-LIFE. IN CELLS OF HEMATOPOIETIC ORIGIN, STRONGLY AND
 CC EXCLUSIVELY DETECTED IN TERMINALLY DIFFERENTIATED MACROPHAGES.
 CC TRANSCRIPTS EXPRESSION APPEARS TO BE DOWN-REGULATED IN PRIMARY
 CC LUNG TUMOR.
 CC -1- INDUCTION: CYTOKINE AND CELLULAR ADHESION TRIGGER FAK INDUCTION.
 CC -1- PFM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
 CC CELLS EXIT MITOSIS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC5/POLO SUBFAMILY.
 CC -1- SIMILARITY: Contains 2 POLO box domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1db-rib.ch/announce/>
 CC or send an email to license@1db-rib.ch).
 CC -----
 CC EMBL: AJ293866; CAC10659.1; -
 CC EMBL: U56998; AAC50637.1; ALT_INIT.
 CC Gene: HGNC:2154; CNK.
 CC Gf: Q9484; -
 CC MIM: 602913; -
 CC GO: GO:0004674; P:protein serine/threonine kinase activity; TAS.
 CC GO: GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC GO: GO:0000074; P:regulation of cell cycle; TAS.
 CC InterPro: IPR000959; POLO box.
 CC InterPro: IPR000719; Prot_Kinase.
 CC InterPro: IPR002290; Ser_Thr_Pkinase.
 CC Pfam: PF00659; Pkinase; 1.
 CC Pfam: PF00659; POLO box; 2.
 CC ProDom: PD000001; Prot_kinase; 1.
 CC SMART: SM00220; S_TKc_1.
 CC DR PROSITE: PSS50078; POLO BOX; 2.
 CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC KM Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
 CC KM Phosphorylation.
 CC FT DOMAIN 62 314 PROTEIN KINASE.
 CC FT NP_BIND 68 76 ATP (BY SIMILARITY).
 CC FT BINDING 91 91 ATP (BY SIMILARITY).
 CC FT ACT_SITE 185 185 BY SIMILARITY.
 CC FT DOMAIN 470 537 POLO BOX 1.
 CC FT DOMAIN 567 637 POLO BOX 2.
 CC FT CONFLICT 99 99 V -> A (IN REF. 2).
 CC FT CONFLICT 353 353 V -> G (IN REF. 2).
 CC FT CONFLICT 419 419 H -> D (IN REF. 2).
 CC FT CONFLICT 470 470 PSBWVG -> VSKMVDY (IN REF. 2).
 CC FT CONFLICT 522 522 R -> P (IN REF. 2).
 CC SQ SEQUENCE 646 AA; 71789 MM; C20147CD0FC8A3B4 CRC64;
 CC Query Match 16.1%; Score 304.5; DB 1; Length 646;
 CC Best Local Similarity 28.9%; Pred. No. 2e-16;
 CC Matches 101; Conservative 48; Mismatches 134; Indels 67; Gaps 15;
 CC
 CC QY RPYOKRAR-----SGPQRLPPCL-----LPLSP-----TAPDRAVATASRL 68
 CC DB 11 RPPORRAAATAAPAGPFP--PFSALRPPBELMAGLTSPPGRLLITPPSGRTYTLKRRLL 68
 CC QY 69 GPVLLLEPEGG--RAYRALHCPGTETGYTKVYPVQBALAVLEPYARLP----- 115
 CC DB 69 G-----KGFARCEATIDTETGSAVAVKVIQSR---VVKPHQREKLINELHRLD 117

QY 116 PHKVAAPTEVLAQTOLLVAFETR-THGDMSLVSRRHRIPEPAALVFROMATALACH 174
 DB 118 QHRIVAFSHHFEDADNIYIFELCSRKSLAHIKWABHTLLEPVRVYRLQILSLGKYLH 177
 QY 175 QHGIVLEDLICRPVFPDRERKULVENLEDSCTLTPDPSLMPKNC--PAYVGPELIS 232
 DB 178 QRGILHMDLKGNEFIT--ENMELKVDGFGLAARLBPPEOR--KRTICGTNYVAPEVIL 233
 QY 233 SRASVSGKADVWSISVALFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLV 292
 DB 234 RQG--KRPADVWSLGGVMTTLGSGPPEFADIKERYRCIKQVHYTLPLASLSPADQL 291
 QY 293 RCLIRREPARLRTATGILLHPWLQ---DPMPLAPFRSLHWEAQQVPD 338
 DB 292 AALIRASPRDRPSIDQLRHDFPKGYTPRLPLFS-----SCVTPD 333
 RESULT 6
 ID SNF1_CANAL STANDARD; PRT; 620 AA.
 AC P52497; 000309;
 DT 01-OCT-1996 (Ref. 34, Created)
 DT 01-NOV-1997 (Ref. 35, Last sequence update)
 DT 30-MAY-2000 (Ref. 39, Last annotation update)
 DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
 GN SNF1.
 OS Candida albicans (Yeast).
 OC Bakariota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxId=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 32354 / B-311;
 RX MEDLINE=98053924; PubMed=9393775;
 RA Peter R., Chang Y.C., Kwon-Chung K.J.;
 RT "A gene homologous to *Saccharomyces cerevisiae* SNF1 appears to be
 RT essential for the viability of *Candida albicans*.";
 RL Infect. Immun. 65:4909-4917(1997).
 RN [2]
 RP SEQUENCE OF 7-620 FROM N.A.
 RC STRAIN=ATCC 32354 / B-311;
 RA Peter R., Kwon-Chung K.J.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
 CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
 CC PROTEIN SNF1. COULD PHOSPHORYLATES CATS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
 CC SIMILARITY). BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1db-rib.ch/announce/>
 CC or send an email to license@1db-rib.ch).
 CC -----
 CC EMBL: U78129; AAB48643.1; -
 CC EMBL: L39263; AAA92456.1; -
 CC DR HSSP: Q63450; 1A06.
 CC DR InterPro: IPR000719; Prot_Kinase.
 CC DR InterPro: IPR002290; Ser_Thr_Pkinase.
 CC Pfam: PF00069; Pkinase; 1.
 CC ProDom: PD000001; Prot_kinase; 1.
 CC SMART: SM00220; S_TKc_1.
 CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC KM Transferase; Serine/threonine-protein kinase; ATP-binding;
 CC phosphorylation; Carbonhydrate metabolism; Nuclear protein.

```
FT DOMAIN 16 29 POLY-HIS.
FT DOMAIN 53 305 PROTEIN KINASE.
FT NP BIND 59 67 ATP (BY SIMILARITY).
FT BINDING 82 82 ATP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
FT MOD_RES 208 208 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 27 8 PO -- AR (IN REF. 2).
FT CONFLICT 228 234 AGPRVDV -> SSVKLMT (IN REF. 2).
FT CONFLICT 242 255 YVMLGRLPFDEF -> GMSCCVVDHSMYSS (IN REF. 2).
FT CONFLICT 387 387 A -> R (IN REF. 2).
FT CONFLICT 416 416 R -> A (IN REF. 2).
FT CONFLICT 494 494 S -> L (IN REF. 2).
SQ SEQUENCE 620 AA; 70005 MW; 18066525061D28 CRC64;

Query Match 16.1%; Score 304; DB 1; Length 620;
Best Local Similarity 29.4%; Pred. No. 2.1e-16;
Matches 95; Conservative 55; Mismatches 133; Indels 42; Gaps 12;

QY 48 LPLSPPTAPDRAVAVATASRLGPLYLLEP-BEG--GRAVYALHCTGTETCKYVQGEA 104
DB 38 VPIDPAPNP-----ANRIGROYQLTGLBGSFGKYLQHLGTQKVALKIIN-RKT 88
QY 105 LAIVE-----PYARLPKHVAVRPTVLAQTQLLYAPFTRTHGDMHSLVSRHNP 155
DB 89 LAKSDMGCRVEREISYLRLLRHPHITIKLYDVYKSKDEIIMVIERFAGKELFDYIVQKGMF 148
QY 156 EPEAAVLFROMATLALHCHQHLVLDLTKCFVPADEBRKKLVLENLEDSCVLTGPPDS 215
DB 149 EDEARRPFOQIIAIVEYCHRHKIYRDLKPEHLLDDQLVNKIDFGL--SNIMT---DG 203
QY 216 LMDKAC--PAYV-GPEILSRBASYSGLADVMSGLVLPMLAGHYFPDSEFVLLFGK 272
DB 204 NFLTSSCSPPYMPAPBEVVISGTL-VAGPEVDVMSGVLVYVLCRLPDPDEFIPLALPK 262
QY 273 IRRGAYALPAGISAPARCLVRCLEBRPAERLTATGILPRLRODPMLPAPTRSHLWEA 332
DB 263 ISNGVYTLPTNLSAGAKILLTRMLVYNPLNRTTHIEHEDMKPKD-MP----- 310
QY 333 AQVVPDGLDLDEAREBEGDREV 355
DB 311 DYLLPPLDLSKNKNSKIDVDDEV 333

RESULT 7
CNK_MOUSE STANDARD; PRT; 631 AA.
ID Q60806; Q60822; Q9R009;
AC 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Cytochrome-inducible serine/threonine-protein kinase (EC 2.7.1.37) (PGF-
DB Inducible kinase).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=NIH Sw188;
RX MEDLINE=95247749; PubMed=7730342;
RA Donohue P.J., Alberts G.F., Guo Y., Winkles J.A.,
RT "Identification by targeted differential display of an immediate early
RL gene encoding a putative serine/threonine kinase.",
RN J. Biol. Chem. 270:10351-10357(1995).
RP SEQUENCE OF 333-437 FROM N.A. (ISOFORM 2).
RC STRAIN=NIH Sw188;
RA Kauejmann G., Weiler M., Kuhl D.,
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RP FUNCTION, AND PHOSPHORYLATION.
```

```
RX MEDLINE=98343954; PubMed=9677325;
RA Chase D., Peng Y., Hanawalt B., Winkles J.A., Longo D.L., Ferris D.K.,
RT "Expression and phosphorylation of fibroblast-growth-factor-inducible
RT kinase (Pnk) during cell-cycle progression.",
RL Biochem. J. 333:655-660(1998).
CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN REGULATING M
CC PHASE FUNCTIONS DURING THE CELL CYCLE. MAY ALSO BE PART OF THE
CC SIGNALING NETWORK CONTROLLING CELLULAR ADHESION. IN VITRO, IS ABLE
CC TO PHOSPHORYLATE CDC25C AND CASIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBUNIT: BINDS TO THE CALCIUM/INTEGRIN-BINDING PROTEIN (CIB). THIS
CC INTERACTION PROBABLY OCCURS VIA THE POLO-BOX DOMAIN (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (by similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q60806-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q60806-2; Sequence=VSP_004927;
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SKIN.
CC -1- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
CC CELLS EXIT MITOSIS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 2 POLO BOX DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 POLO BOX DOMAINS.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as the content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U21392; AAC52191.1; -.
CC EMBL; U22434; AAC52192.1; -.
CC EMBL; AF136586; AAF08369.1; -.
CC PIR; A57286; A57286.
CC MGI; MGI:109604; Cnk.
CC InterPro; IPR000959; POLO box.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_Thr_kinase.
CC Pfam; PF00063; Kinase_1.
CC Pfam; PF00659; POLO_box; 2.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00076; POLO_BOX; 2.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 63 315
FT NP BIND 69 77 ATP (BY SIMILARITY).
FT BINDING 92 92 ATP (BY SIMILARITY).
FT ACT_SITE 186 186 BY SIMILARITY.
FT DOMAIN 455 518 POLO BOX 1.
FT DOMAIN 552 622 POLO BOX 2.
FT VASPLIC 373 373 L -> LVSGMRTSGIHPDVAPB (in isoform 2).
FT CONFLICT 386 386 /FTId=VSP_004927.
FT CONFLICT 386 386 I -> V (IN REF. 2).
SQ SEQUENCE 631 AA; 70012 MW; 20857341870DB102 CRC64;

Query Match 15.8%; Score 298.5; DB 1; Length 631;
Best Local Similarity 28.3%; Pred. No. 5.8e-16;
Matches 98; Conservative 48; Mismatches 133; Indels 67; Gaps 14;

QY 39 PQLPLPCLPLSPPTAPDRAVAVATAS-RIGPLYLLEP----- 77
DB 10 PRP-FPRAAVSAPAPAPG--PPANASPRSEPEVLGAPRAPDPGRLLITDPLSGRTYTK 65
QY 78 -----EGG--RAYVALHCTGTETCKYVQGEALVLEPYARLP-----PRKH 119
```

```

Db      66 GRLLKGGFACVATDTRESGIAVAVKIPQSR---VAKPQREKILNEIHLADLQHRH 122
      120 VAPTEVLACTQLLYAFETR--THGDMHSLVSRHRIPEPAVALFROMATALAHCHQHG 178
      123 IVRSHHPEADNIIYIFLELCSRKSLAHIMKARHTLEPEVRYLLQSLGKYLHQGCI 182
      179 VLRLKLCRFVPADEKKVLLENLEDSCVLTGPDLSMDKAC--PAYVPELISRSAS 236
      183 LHRDLKLGKGFPTD--NMELKVGDFGLAARLEPEQR--KKTICGTPNYVAPVLLRQG- 237
      237 YSGKADVMSLGVALLFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRL 296
      238 -HGEADVMSLGCVMYTLCSPPERTADLKSTRICQKVHYTLPSLSLPAKQLLAAIL 296
      297 RREPAERLTATGILLHPMLRQ---DPMPLAPTRSHLMEAAQVVPD 338
      297 RASPRDRPSIEQLRHDFFTKGYTPDRLPVS-----SCVTVPD 334

```

RESULT 8

AC 09R011; ID CNK_RAT STANDARD; PRT; 615 AA.

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (PGF-
DE Inducible kinase) (fragment).
GN CNK OR FNK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=994523297; PubMed=10523297;
RA Kauselmann G., Weiler M., Wulff P., Jesberger S., Konietzko U.,
RA Seafeldt J., Staubli U., Bereliter-Hahn J., Strechardt K., Kuhl D.,
RT "The polo-like protein kinases Pmk and Snk associate with a Ca(2+)-and
RT integrin-binding protein and are regulated dynamically with synaptic
RT plasticity."
RL EMBL J. 18:5528-5539(1999).
CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN REGULATING M
CC PHASE FUNCTIONS DURING THE CELL CYCLE. MAY ALSO BE PART OF THE
CC SIGNALING NETWORK CONTROLLING CELLULAR ADHESION. IN VITRO, IS ABLE
CC TO PHOSPHORYLATE CDC25C AND CASBIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBUNIT: BINDS TO THE CALCINUM/INTEGRIN-BINDING PROTEIN (CIB). THIS
CC INTERACTION PROBABLY OCCURS VIA THE POLO-BOX DOMAIN
CC -1- SUBCELLULAR LOCATION: WHEN INDUCED, IT TRANSLOCATES INTO THE
CC DENDRITES OF ACTIVATED NEURONS.
CC -1- TISSUE SPECIFICITY: CONSTITUTELY EXPRESSED IN POST-MITOTIC
CC NEURONS.
CC -1- INDUCTION: BY THE INTENSE ACTIVITY ASSOCIATED WITH SEIZURES.
CC -1- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
CC CELLS EXIT MITOSIS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC5/POLO SUBFAMILY.
CC -1- SIMILARITY: Contains 2 POLO box domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL, AF136584; AF08367.1; -
DR InterPro: IPR000959; POLO_box.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.

```

```

DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00659; POLO_box; 2.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS50079; POLO_BOX; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW phosphorylation.
FT NON TER 1 1
FT DOMAIN 31 283 PROTEIN KINASE.
FT BP BIND 37 45 ATP (BY SIMILARITY).
FT BINDING 60 60 ATP (BY SIMILARITY).
FT ACT_SITE 154 154 BY SIMILARITY.
FT DOMAIN 440 503 POLO BOX 1.
FT DOMAIN 537 607 POLO BOX 2.
FT NON TER 615 615
SQ SEQUENCE 615 AA; 68800 MW; 06584C229B5D71BD CRC64;

```

Query Match 15.4%; Score 291; DB 1; Length 615;
Best Local Similarity 28.5%; Pred. No. 2.2e-15;
Matches 94; Conservative 47; Mismatches 127; Indels 62; Gaps 13;

```

Cy 41 PRLPCLLPSPRPAPDRATAVATASRLGCVLLLEPEEG-----GRAYAL 86
Db 3 PRSEPEV--LAGFPTD-----ASRL-----ITDPSGRTYIKRLLGKGFACVYAT 49
Cy 87 HCPGTGTCVKYVPGALAVLEPYARLP-----PHKVARPTVLAQTQLYA 135
Db 50 DTEISIVAVKVIPOSr---VAKPHQREKILNEIHLADLQHRIVFSSHPEADNIIYI 106
Cy 136 FPTR--THGDMHSLVSRHRIPEPAVALFROMATALAHCHQGLVLDLKLGRVPADE 194
Db 107 FLRLCSRKSLAHIMKARHTLEPEVRYLLQSLGKYLHQRGITLHRDLKLGKGFPTD-- 164
Cy 195 RKGLVLENLEDSCVLTGPDLSMDKAC--PAYVPELISRSASVSGKADVMSLGVALL 252
Db 165 NMELKVGDFGLAARLEPEQR--KKTICGTPNYVAPVLLRQG--HGEADVMSLGCVMY 220
Cy 253 TMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRLRREPAERLTATGILLH 312
Db 221 TLGCGSPPERTADLKSTRICQKVHYTLPSLSLPAKQLLAAILRASPRDRPSIEQLRH 280
Cy 313 PWLRQ---DPMPLAPTRSHLMEAAQVVPD 338
Db 281 DPFKGYTPDRLPVS-----SCVTVPD 302

```

RESULT 9

AC P34244; ID KKK1_YEAST STANDARD; PRT; 1518 AA.

```

DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.-).
GN YKL101W OR YKL453.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=8288C;
RX MEDLINE=94078677; PubMed=8256524;
RA Pallier C., Valens M., Puzos V., Fukuhara H., Charet G., Sor P.,
RA Bolestin-Fukuhara M.,
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
RT physically localizes the MRL1 gene and reveals eight new open reading
RT frames, including a homologue of the KIN1/KIN2 and SNF1 protein
RT kinases."
RL Yeast 9:1149-1155(1993).

```


CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIN1 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.iesb-sib.ch/announce/>
 CC or send an email to license@iesb-sib.ch).
 CC -----
 CC EMBL: X71133; CAA50456.1; -
 CC EMBL: Z28101; CAA81941.1; -
 CC PIR: S37928; S37928.
 CC HSSP: Q63450; 1A06.
 CC SCD: S0001584; HSL1.
 CC GO: GO:0005935; C:bud neck; IDA.
 CC GO: GO:0005940; C:septin ring; IDA.
 CC GO: GO:0000086; P:G2/M transition of mitotic cell cycle; IGI.
 CC GO: GO:0006468; P:protein amino acid phosphorylation; IDA.
 CC GO: GO:0000074; P:regulation of cell cycle; IMP.
 CC InterPro: IPR000719; Prot. kinase.
 CC InterPro: IPR002290; Ser. thr. kinase.
 CC InterPro: IPR001245; Tyr. kinase.
 CC Pfam: PF00069; pkinase; 1.
 CC PRINTS: PR00109; TYRKINASE.
 CC ProDom: PD000001; Prot. kinase; 1.
 CC SMART: SMO0220; S_TKC; 1.
 CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE: PS00108; PROTEIN KINASE ST; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC Hypothetical protein; transferase; Serine/threonine-protein kinase;
 CC ATP-binding.
 CC DOMAIN 81 369 PROTEIN KINASE.
 CC NP_BIND 87 95 ATP (BY SIMILARITY).
 CC BINDING 110 110 ATP (BY SIMILARITY).
 CC ACT_SITE 239 239 BY SIMILARITY.
 CC SEQUENCE 1518 AA; 169592 MW; 803f84f531241DD CRC64;
 SQ
 Query Match 15.2%; Score 287; DB 1; Length 1518;
 Best Local Similarity 26.7%; Pred. No. 1.3e-14;
 Matches 86; Conservative 50; Mismatches 126; Indels 58; Gaps 9;
 QY 56 PDRAVAVTAS-----RLGRVYL---LEPEGRAVALHCPGTETTCVYVQAL 105
 DB 59 PDSTVAVATKSKRSKRSRDVGPWKLKGLTGKSGSRVALAKMETGOLAIKIVKKA 118
 QY 106 -----AVLEPYA-----RLPRKHVAR 122
 DB 119 VHCNMGVTPNSYSSMTSNVSSPSIASREHSHSQTNPFIETREIYIMKLISHTNWA 178
 QY 123 PTEVLATQLLYAFRTTH-GDMSLVSRRIIPPEAAVLFRQATALAHCHQGLVLR 181
 DB 179 LPEVVENKSELVLEVDYDGLFPIYLVSKGLPEREAIHYKQIVGEVSYCHSNICHR 238
 QY 182 DLKCRFPADREKRLVLENIEDSCVLTGPDSDMDKACAYVGPRIISRSASYSKA 241
 DB 239 DLK-PENLIDKKNRIKIDP-GMAALELPKRLKTSGSPHYASPEIWMGR-PYHGGP 295
 QY 242 ADVASLVALFTMLAGHYPODSBPVLLFGKIRGAYVALPAGLSAPACLVACLRSPA 301
 DB 296 SDVMSGCVLFPALLTGLHPFNDNDIKKLLKVKOSKTYQMPNLSSEADLISKIIVDPE 355
 QY 302 ERLTATGILLHPWLRQ-DPMPL 322
 DB 356 KRITTOEILKHPILIKKYDDLVP 377

DT 01-NOV-1997 (Rel. 35, Last Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 36, Last annotation update)
 DB Carbon catabolite derepressing protein kinase (EC 2.7.1.1-).
 GN SNP1.
 OS Candida glabrata (Yeast) (Torulopsis glabrata).
 OC Saccharomycetales; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OK NCBI_TaxID=5478;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCL584;
 RX MEDLINE=97101049; PubMed=8945576;
 RA Peller R., Kwon-Chung K.J.;
 RT "Disruption of the SNF1 gene abolishes trehalose utilization in the
 RT pathogenic yeast Candida glabrata.";
 RL Infect. Immun. 64:5269-5273(1996).
 CC -1- FUNCTION: ESSENTIAL FOR GLUCOSE REPRESSION. IT
 CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
 CC PROTEIN SNF4, COULD PHOSPHORYLATES CAT8 (BY SIMILARITY)
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIN1 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.iesb-sib.ch/announce/>
 CC or send an email to license@iesb-sib.ch).
 CC -----
 CC EMBL: L78130; AAB48642.1; -
 CC HSSP: P24941; 1HCL.
 CC InterPro: IPR000719; Prot. kinase.
 CC InterPro: IPR002290; Ser. thr. kinase.
 CC Pfam: PF00069; pkinase; 1.
 CC ProDom: PD000001; Prot. kinase; 1.
 CC SMART: SMO0220; S_TKC; 1.
 CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE: PS00108; PROTEIN KINASE ST; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC Transferase; Serine/threonine-protein kinase; ATP-binding;
 CC Phosphorylation; Carbohydrate metabolism; Nuclear protein.
 CC DOMAIN 6 17 POLY-HIS.
 CC NP_BIND 39 290 PROTEIN KINASE.
 CC BINDING 45 53 ATP (BY SIMILARITY).
 CC ACT_SITE 68 68 ATP (BY SIMILARITY).
 CC MOD_RES 161 161 BY SIMILARITY.
 CC SEQUENCE 611 AA; 70049 MW; 89E17812A4900CDO CRC64;
 SQ
 Query Match 15.0%; Score 284.5; DB 1; Length 611;
 Best Local Similarity 27.9%; Pred. No. 7e-15;
 Matches 90; Conservative 54; Mismatches 136; Indels 43; Gaps 10;
 QY 57 DRATAVATASRLGVLLEP-EEG--GRAYRALHCPGTETTCVYVQALAVLE----- 109
 DB 25 NKVSLADGSRVGVYQVLTGEGSPFGKVLAVHTTQKVALKINSIN-KVYLAQSDMGOR 83
 QY 110 -----PYAKRLPHGHVARPTVLAQTQLLYAFRTTHGDMSLVSRRIIPPEAAVLFR 164
 DB 84 IDREISYRLRLRHPIIKLYDVINSKDEIIVWIEYAGNELDYVQRKMSBOEARPFQ 143
 QY 165 QMATALAHCHQGLVLDLKLKRFVADREKRLVLENIEDSCVLTGPDSDMDKAC-- 222
 DB 144 QIISAVKCHRHKIVHDLKRENLLDBENLVKADTGL--SNMT--DGNFLKTSGS 198
 QY 223 PAYVGPRIISRSASYSKADVMSLVALFTMLAGHYPODSBPVLLFGKIRGAYVALPA 282
 DB 199 PNYAPAPVISGKL-YADPEVDVMSGCVILYVMLCRRLPFDSEIPVLFKNISNGVYTLPK 257


```

OY 283 GLSAPARCLVRCILRRREPRLTATGILHPWLKQDPMPLAPTRSHLMEAAQVYPDL-- 340
DB 258 FLSPGASDLIKRMLIVPLNKRISLHEKMODEWFRVD-----LABYVPODLKQ 305
OY 341 -----GLDEAREEEDREVV 355
DB 306 QBOFNKSGNBNENVEIDENVV 328

RESULT 11
K110 ARATH STANDARD; PRT; 512 AA.
ID K110 ARATH 004728; Q39076.
AC 038997; 004728; Q39076.
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE SNF1-related protein kinase KIN10 (EC 2.7.1.-) (AKIN10).
GN KIN10 OR SKIN10 OR AT3G01090 OR TAP13.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702.
RN 1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93013041; Pubmed=1339373;
RA 1e Guen L., Thomas M., Blanchi M., Halford N.G., Kreis M.;
RT "Structure and expression of a gene from Arabidopsis thaliana
RL encoding a protein related to SNF1 protein kinase.";
RN 120:249-254(1992).
RN 12]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RA Lessard P., Kreis M.;
RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN 13]
RP SEQUENCE FROM N.A.
RX MEDLINE=21016720; Pubmed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unseld M.,
RA Faltmann B., Valle G., Bloeker H., Perez-Alonso M., Obermayer B.,
RA Delany M., Boutry M., Givell L.A., Maché R., Pulidomenech P.,
RA De Simone V., Cholane N., Artiguenave F., Robert C., Brothier P.,
RA Wincker P., Cartolico L., Weisenbach J., Saurin W., Querlet F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmach E., Drzonek H., Britle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Vose H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simonati B.,
RA Conrad A., Horstschler K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Scharte M., Schoen O., Bagnues M., Terol J., Clément J.,
RA Navarro P., Collado C., Perez-Perez A., Ottensmelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Delauzo C., Purnelle B., Maury D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casasuberta E.,
RA Monfort A., Argilou A., Flores M., Liguori R., Vitale D.,
RA Manhaupt G., Haase D., Schoof H., Ruid S., Zaccaria P., Mewes H.-W.,
RA Meyer K.P.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T.H., Rizzo B., Walts A., Utecherback T., Fujii C.Y., Shea T.P.,
RA Creeby T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Miltchev J., Seilers P., Gill J.B., Feldblyum T.V.,
RA Preuss D., Lin X., Niermen W.C., Salzberg S.L., White O., Venter J.C.,
RA Frazer C.M., Kaneko T., Nakamura K., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Ideawara Y., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RN Nature 408:820-822(2000).
RN 14]
RP SEQUENCE OF 1-19 FROM N.A.
RX STRAIN=cv. Columbia;
RX MEDLINE=95115691; Pubmed=7816049;
RA 1e Guen L., Thomas M., Kreis M.;

```

```

RT "Gene density and organization in a small region of the Arabidopsis
RT thaliana genome.";
RL Mol. Gen. Genet. 245:390-396(1994).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN A SIGNAL TRANSDUCTION
CC CASCADE REGULATING GENE EXPRESSION AND CARBOHYDRATE METABOLISM IN
CC HIGHER PLANTS
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ROOTS, SHOOTS AND LEAVES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC -1- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ch/announce/
CC or send an email to license@ebi.ebi.ch).
CC -----
DR EMBL; M93023; AAA3736.1; -
DR EMBL; X79707; CAA56146.1; -
DR EMBL; AC008261; AAF26165.1; -
DR PIR; J14446; J14446.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR001772; Kinase_Cterm.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF02149; KA1; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50030; UBA; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 19 271 PROTEIN KINASE.
FT NP_BIND 292 332 UBA.
FT BINDING 25 33 ATP (BY SIMILARITY).
FT ACT_SITE 48 48 ATP (BY SIMILARITY).
FT MOD_RES 142 142 BY SIMILARITY.
FT MOD_RES 175 175 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 512 AA; 58373 MW; 5A18655A0A506DF CRC64;

Query Match 14.9%; Score 281; DB 1; Length 512;
Best Local Similarity 28.7%; Pred. No. 1,1e-14;
Matches 87; Conservative 46; Mismatches 118; Indels 52; Gaps 9;

OY 80 GRAYRALMCPSTYETCYKY-----PVGEALVLBRVAPLPKHVAPRTYVL 127
DB 31 GAVKIAEHALTGKVAIKILNRKIKNMEMBEKVREBKIL-----RLPHNPHIRLYEVI 86
OY 128 AGTQLLYAFPTTH-GDMSLYRSRRRIEPEPAVLFRQMAATLAAHCHQGLVLRDLK- 184
DB 87 ETPFDLYLWMEYVNGELDYIVKGRLOEDERARNFQOIIIGVEYCHRMVVHRDLKPE 146
OY 185 -----LCRFVADREKKLVLEN--LSDCVLTGPDSDSLMDKACPAVVGPEILSSPAS 236
DB 147 NLLDSKCNVKIADGLGSLNMDGHPKLTSC-----GSPNYAAAEVIGSKL- 192
OY 237 YSGKADWSLGVALFTMLAGHYPFQDSPEVLLFGKIRGAVALPAGLSAPARCLVRCIL 296
DB 193 YAGPEVDWSCVILYALICGLTPDDEENIPMLFKIKIGIYTLPSHLSPGARDLIPRL 252
OY 297 REPARBRLTANGILLHPWLKQDPMPLAPTRSHLMEAAQV--VDDGLGLDAREEEDREVV 355

```

DB 253 VQDPMKRVTEIRQHMPF-----QAHLPYLAVPPTVQAKKIDETILOEVI 302

OY 356 LYG 358

DB 303 NMG 305

RESULT 12

SNF1_YEAST STANDARD; PRT; 633 AA.

ID _SNF1_YEAST

AC P06782

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Carbon catabolite depressing protein kinase (EC 2.7.1.-).

GN SNF1 OR CARI OR CCR1 OR PAS14 OR GLC3 OR YDR477W OR D8035.20.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OC NCBI_TaxId=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86289463; PubMed=3526554;

RA Celens J.L., Carlson M.,

RT "A yeast gene that is essential for release from glucose repression

RT encodes a protein kinase."

RL Science 233:1175-1180(1986).

RN [2]

RP SEQUENCE FROM N.A.

RA Dietrich P.S., Mulligan J., Allen E., Araujo R., Aviles E.,

RA Berno A., Carpenter J., Chen B., Cherry J.M., Chung B., Duncan M.,

RA Huntire-Smith S., Hyman R., Komp C., Lahekard D., Lew H., Lin D.,

RA Moseedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,

RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,

RA Winant A., Yelton M., Botstein D., Davis R.W.,

RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 274-284, 528-539 AND 622-630, AND PHOSPHORYLATION SITE.

RX MEDLINE=94131988; PubMed=7905477;

RA Mitchell K.I., Stapleton D., Gao G., House C., Mitchell B.,

RA Karsela P., Witters L.A., Kemp B.E.,

RT "Mammalian AMP-activated protein kinase shares structural and

RT functional homology with the catalytic domain of yeast Snf1 protein

RT kinase."

RL J. Biol. Chem. 269:2361-2364(1994).

CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT

CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY

CC PROTEIN SNF4. INTERACTS ALSO WITH S1P1, S1P2 AND GAL83. COULD

CC PHOSPHORYLATES CAT8.

CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE.

CC -1- SIMILARITY: BELONGS TO THE SBR/THR FAMILY OF PROTEIN KINASES.

CC SNF1 SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>

CC or send an email to license@ebi.ac.uk).

CC -----

DR EMBL M13971 AAA35058.1; -

DR EMBL U33050 AAB64904.1; -

DR PIR A26030 A26030.

DR HSRP P24941 1HCL.

DR SGD S000285 SNF1.

DR GO GO:0005737 C:cytoplasm; IPI.

DR GO GO:0005634 C:nucleus; IPI.

DR GO GO:0000324 C:vacuole (yeast Fungi); IPI.

DR GO GO:0004679 F:SNF1A/AMP-activated protein kinase activity; IDA.

DR GO GO:0006468 P:protein amino acid phosphorylation; IDA.

DR GO GO:0006995 P:response to nitrogen starvation; IDA.

DR InterPro: IPR000719, Prot_kinase.

DR InterPro: IPR002290; Ser_thr_kinase.

DR Pfam: PF00069; kinase; 1.

DR ProDom: PD000001; Prot_kinase; 1.

DR SMART: SM00220; S_TKc; 1.

DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE: PS00108; PROTEIN KINASE ST; 1.

DR PROSITE: PS00011; PROTEIN KINASE_DOM; 1.

DR Transferase; Serine/threonine-protein kinase; ATP-binding;

DM Phosphorylation; Carbohydrate metabolism; Nuclear protein.

KW DOMAIN 18 32 POLY-HIS.

FT NP BIND 55 306 PROTEIN KINASE.

FT NP BIND 61 69 ATP (BY SIMILARITY).

FT BINDING 84 84 ATP (BY SIMILARITY).

FT ACT SITE 177 177 BY SIMILARITY.

FT MOD RES 210 210 PHOSPHORYLATION (AUTO-).

SO SEQUENCE 633 AA; 72045 MW; F5C63565C986C483 CRC64;

Query Match 14.5%; Score 274.5; DB 1; Length 633;

Best Local Similarity 27.8%; Pred. No. 4,4e-14;

Matches 88; Conservative 53; Mismatches 138; Indels 37; Gaps 11;

OY 51 SPTAPDRATVATVATAGPYVLLRP--EAG--GRAVRLHCPGTBYTCRVYVQBALAV 107

DB 35 SNETLNNPKSLADGAGHNGYQVKTIGSGSPKVKLAHYTTGQKVALKIN-KYVLAK 93

OY 108 LE-----PYRLPEKHVARPTEVTLACTQLLVAFFRTQDMHSLVSRHRIRPEPR 158

DB 94 SDMGRIEREISYLRILRHPIIKLYDVVSKDEIIVIEYAGNELPDYVQORMSBOE 153

OY 159 AAVLFRQMAALAHCHQHGVLDLKLRFVADRERKKVLBNLBSCVLTGPDSDLMD 218

DB 154 ARRPFOQIISAVSECHHKVHRDLLENILDSHLNKLADGL--SNIT--DGNFL 208

OY 219 KHAC--PAVYGPRLISRSASGKAADVMSLGVALLPTMLAGHYPFQDSEPTLLFGKIRG 276

DB 209 KTCGSGNVAPAEVIGSKL-YAGPVDVWGCYILVYMLGRRLPFDBESIPVLFKINSNG 267

OY 277 AYALPAGLSAPARCLVRCILRRPABERLITATGILHPWLROD-PMPLAPRSHLMEAAQV 335

DB 268 VYTLPEKPLSGAAGLIRKMLIVNPLNRSIHEIMQDWFVVDLPVEYLLP----- 316

OY 336 VPDGLGLDEAREEGD 351

DB 317 PD---LKPHEBENE 328

RESULT 13

KCCD_RAT STANDARD; PRT; 533 AA.

ID _KCCD_RAT

AC P15791

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Calcium/calmodulin-dependent protein kinase type II delta chain

DE (EC 2.7.1.123) (Cam-kinase II delta chain) (Cam kinase II delta

DE subunit) (CaMK-II delta subunit).

GN CAMK2D.

OS Rattus norvegicus (Rat).

OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Eumetazoa; Butleria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC NCBI_TaxId=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90036861; PubMed=2553697;

RA Tobimatsu T., Fujisawa H.,

RT "Tissue-specific expression of four types of rat calmodulin-dependent

RT protein kinase II mRNAs."

RL J. Biol. Chem. 264:17907-17912(1989).

RN [2]

RP SEQUENCE OF 314-533 FROM N.A.

RX STRAIN=Sprague-Dawley; TISSUE=Aorta, and Skeletal muscle;

RX MEDLINE=9330844; PubMed=8390994;

RA Schworer C.M., Rothblum L.I., Thekkumkara T.J., Singer H.A.;

```

RT      "identification of novel isoforms of the delta subunit of
RT      Ca2+/calmodulin-dependent protein kinase II. Differential expression
RT      in rat brain and aorta."
RL      J. Biol. Chem. 268:14443-14449(1993).
CC      -1- FUNCTION: THIS KINASE MAY PLAY A ROLE IN NEUROTRANSMISSION.
CC      -1- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.
CC      -1- ENZYME REGULATION: AUTOPHOSPHORYLATION OF CAM-KINASE II PLAYS AN
CC      IMPORTANT ROLE IN THE REGULATION OF THE KINASE ACTIVITY.
CC      -1- SUBUNIT: COMPOSED OF FOUR DIFFERENT CHAINS: ALPHA, BETA, GAMMA,
CC      AND DELTA.
CC      -1- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named Isoforms=4;
CC      Name=Delta 1;
CC      IsoId=PI5791-1; Sequence=Displayed;
CC      Name=Delta 2;
CC      IsoId=PI5791-2; Sequence=VSP_004784;
CC      Name=Delta 3;
CC      IsoId=PI5791-3; Sequence=VSP_004785, VSP_004786, VSP_004787;
CC      Name=Delta 4;
CC      IsoId=PI5791-4; Sequence=VSP_004788;
CC      -1- TISSUE SPECIFICITY: DELTA 1 IS THE PREDOMINANT FORM IN THE BRAIN,
CC      DELTA 2 AND 3 PREDOMINATE IN THE AORTA AND DELTA 4 IN SKELETAL
CC      MUSCLE.
CC      -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. CAMK
CC      SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL collaboration -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi-sdb.ch/announce/
CC      or send an email to license@ebi-sdb.ch).
CC      -----
DR      EMBL, J05072; AAA40866.1; -
DR      EMBL, U13406; AAA41479.1; -
DR      EMBL, U13407; AAA41480.1; -
DR      EMBL, U13408; AAA41481.1; -
DR      PIR, A34366; A34366.
DR      HSSP, O63450; 1A06.
DR      InterPro, IPR000279; Prot_kinase.
DR      InterPro, IPR002290; Ser_thr_kinase.
DR      Pfam, PF00069; pkinase; 1.
DR      ProDom, PD00001; Prot_kinase; 1.
DR      SMART, SM00220; S_TKc; 1.
DR      PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE, PS00108; PROTEIN_KINASE_ST; 1.
DR      PROSITE, PS50011; PROTEIN_KINASE_DOM; 1.
KW      Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW      phosphorylation; ATP-binding; Neurone; Alternative splicing.
FT      DOMAIN 14 272
FT      NP_BIND 10 28
FT      BINDING 43 43
FT      ACT_SITE 136 136
FT      DOMAIN 291 301
FT      VARSPLIC 329 362
FT      FTID=VSP_004784.
FT      VARSPLIC 329 335
FT      FTID=VSP_004785.
FT      VARSPLIC 337 359
FT      FTID=VSP_004786.
FT      VARSPLIC 360 362
FT      FTID=VSP_004787.
FT      VARSPLIC 349 362
FT      FTID=VSP_004788.
FT      SEQUENCE 533 AA; 60080 MW; B41BCB2BSA0DB7CA CRC64;
Query Match 14.5%; Score 273.5; DB 1; Length 533;
Best Local Similarity 26.8%; Pred. No. 4.3e-14;
Matches 79; Conservative 53; Mismatches 123; Indels 19; Gaps 7;
61 AVATRSRLGPFVLLPEEGGAY---RALHCPTGTETCYVYQDEALA---VLEPYA 112
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Dd		STTTCTGTPTDEYGLFBEELGKGAFSVVRCKMIFPGQGYAAKIINTTKLSARDHQIKLERRA	62
Oy	3	---	
Oy	113	--RFLPHKHAVAPRTVLATGQLLYAFTR--THDMSIVSRNRIRIPEPAVALFRGMAT	168
Dd	63	RICLLKRPNINVRLDHSDISESGFYLVFDLTVGSELFEEDIVARYRYSEADASHICTOILE	122
Oy	169	ALARCHOHGVLRLDLKCRFPVFADRERKKLV-LENLBDSCVLTGPDDSLMDHKACPAYGV	227
Dd	123	SVMCHLANGVIHRRLKEKNILLASKSKGAAYKLADPCLAIHYGGDDQAAMGFACTPCYLS	182
Oy	228	PETISSRASYSKGADWVSIGVALEFTMLAGHYPPQSSPVALLFGKITRGAYVALPA---G	283
Dd	183	PERVL-RNDPYGRKVDMMACGVILYYILVGYPPEPWDEDQRHLTQQIQAKAGAVDPSPENDT	240
Oy	284	LSAPARCLVRCLEARREPARLTAHGILHHMWLRQ	317
Dd	241	VTPRAKDILINKMTLINPKAKRIYASEALKHPMICQ	274
<hr/>			
	RESULT 14		
ID	CDRI_SCHPO	STANDARD;	PRT; 593 AA.
AC	P07J34; Q9P6Q4;		
DT	01-APR-1988 (Rel. 07, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Mitosis inducer protein kinase cdri (EC 2.7.1.-) (Protein kinase nhl1).		
Gn	CDRI OR NIM1 OR SPAC644.06C.		
Os	Schizosaccharomyces pombe (Fission yeast).		
Oc	Bukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
Oc	Schizosaccharomycetales; Schizosaccharomycetaceae;		
Ox	Schizosaccharomycetes.		
RX	NCBI_TaxID=4896;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91169281; PubMed=2004705;		
RA	Fellotter H., Nurse P., Young P.G.;		
RT	"Genetic and molecular analysis of cdri/niml in Schizosaccharomycetes		
RT	pombe.";		
RL	Genetics 127:309-318(1991).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=87187654; PubMed=3453113;		
RA	Russell P., Nurse P.;		
RT	"The mitotic inducer niml+ functions in a regulatory network of		
RT	protein kinase homologs controlling the initiation of mitosis.";		
RL	Cell 49:569-576(1987).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=972;		
RX	MEDLINE=21846401; Pubmed=11859360;		
RA	Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,		
RA	Sources J., Peat N., Hayles J., Baker S., Baham D., Bowman S.,		
RA	Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,		
RA	Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,		
RA	Gentile S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,		
RA	Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,		
RA	Jones K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,		
RA	Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,		
RA	Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,		
RA	Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,		
RA	Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,		
RA	Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,		
RA	Woodward J., Voicakeert G., Aert R., Robben J., Gymnopreux B.,		
RA	Weljens I., Vanstreets S., Rieger M., Schaefner M., Mueller-Auer S.,		
RA	Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,		
RA	Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,		
RA	Eger F., Zimmermann W., Medler U., Wandut R., Furnelle B.,		
RA	Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,		
RA	Gallibert F., Ayas S.U., Xiang Z., Hunt C., Moore K., Hurst S.M.,		
RA	Lucas M., Rochee M., Galliardin C., Tallada V.A., Garzon A., Thode G.,		
RA	Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito T.,		

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Usery D., Barrett B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: THIS PROTEIN, A DORS-DEPENDENT MITOTIC INDUCER, APPEARS
 CC TO FUNCTION AS A NEGATIVE REGULATOR OF MITOSIS INHIBITOR WEI1 BY
 CC PHOSPHORYLATING AND INACTIVATING IT.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIM1 SUBFAMILY.
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL, X57549; CAA40774.1; -
 DR EMBL, AL355012; CAB90133.1; -
 DR EMBL, AL355012; CAB90133.1; -
 DR PIR, S16153; KIZPMN.
 DR HSSP, Q63450; 1A06.
 DR GenedB_Spomb; SPAC644.06c; -
 DR InterPro; IPR000719; Prot. Kinase.
 DR InterPro; IPR002290; Ser. Thr. Kinase.
 DR InterPro; IPR01345; Tyr. Kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot. Kinase; 1.
 DR SMART; SMO0220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KM MitoS16.
 FT DOMAIN 12 258 PROTEIN KINASE.
 FT NP_BIND 18 26 ATP (BY SIMILARITY).
 FT BINDING 41 41 ATP (BY SIMILARITY).
 FT ACT_SITE 128 128 BY SIMILARITY.
 FT CONFLICT 252 252 V -> F (IN REF. 1).
 FT CONFLICT 570 570 A -> I (IN REF. 1).
 FT SEQUENCE 593 AA; 66954 MW; DE30AE06B070F458 CRC64;
 SQ
 Query Match 14.4%; Score 271.5; DB 1; Length 593;
 Best Local Similarity 31.5%; Pred. No. 7e-14;
 Matches 74; Conservative 45; Mismatches 107; Indels 9; Gaps 6;
 Oy 85 ALHCEPTGTEYCKVYVQVDAVALE-PYARLPKHKVAPPEVLAGTOLLY-AFPTRTHG 142
 Db 29 AKHAKTGLAIAIKIPRYASGMEILMMRLRHNIILYDVMTDQHMVYALTYVDDG 88
 Oy 143 DMHSLVRRNRHIREPEAAVLPROMATALAHCHQHLVLRDLKCFVADDERKRLVLEN 202
 Db 89 ELFYIRKRGPSERSEAAHYLSQIIDAVAHCHRPFRHRLDLLENILIKVNEQQ--IKI 145
 Oy 203 LEDSCVLTPDSDSLMDKACPA--YVGEIILSRASYSKADAVSLGVALLFTMLAGHP 260
 Db 146 ADPGMAVTEPRNDSCLENV-CGSLHYLAPRIVSHK-PYNGAPADVSCGVILSLSNLCP 203
 Oy 261 FDSSEPVLLFGKIRIGAYALPAGLSAPARCLVRCLRRBPARKLTATGILLHPWL 315
 Db 204 FGQGNVTIVYKIRIGAYDLPSSISAAODLHRMLDVNPSTRTITPVFSSHPL 258
 RESULT 15
 SNP1_SCHPO STANDARD; PRT; 576 AA.
 ID SNP1_SCHPO
 AC 074536;
 DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DB SNP1-11ke protein kinase (EC 2.7.1.-).
 GN SPCC74.03c.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21849401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornby I., Howarth S., Huckle E.J., Hunt S., Jagsle K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Mablet D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinulitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Roben J., Grymoprez B.,
 RA Welljens I., Vansteele E., Rieger M., Schefer M., Mueller-Auer S.,
 RA Gabel K., Fuchs M., Fritzc C., Holzer E., Moesli D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Laureau V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez J., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Usery D., Barrett B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIM1 SUBFAMILY.
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL, AL031543; CAA20833.1; -
 DR PIR, T41587; T41587.
 DR HSSP, Q63450; 1A06.
 DR GenedB_Spomb; SPCC74.03c; -
 DR InterPro; IPR000719; Prot. Kinase.
 DR InterPro; IPR002290; Ser. Thr. Kinase.
 DR InterPro; IPR000449; UBA domain.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot. Kinase; 1.
 DR SMART; SMO0220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50030; UBA; 1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KM DOMAIN 34 285 PROTEIN KINASE.
 FT DOMAIN 304 345 UBA.
 FT NP_BIND 40 48 ATP (BY SIMILARITY).
 FT BINDING 63 63 ATP (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 FT SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;
 SQ

GenCore version 5.1.6
Copyright: (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:55:46

(without alignments)
1721.419 Million cell updates/sec

Title: US-09-909-474D-2

Perfect score: 1891

Sequence: 1 MRATPLAARAGSLSRKKRLR.....GLGIDBARBEGDREVLYG 358

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: PIR.76:.*
2: PIR1:.*
3: PIR2:.*
4: PIR3:.*
5: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	298.5	15.8	631	2 A57286	probable serine/th
2	298	15.8	887	2 T20941	hypothetical prote
3	292.5	15.5	651	2 S52244	p595g3 protein - A
4	291	15.4	602	2 S72513	FG2 protein - yea
5	287	15.2	1518	2 S37928	probable purine nu
6	284	15.0	504	2 T10449	probable serine/th
7	284	15.0	512	2 T52633	serine/threonine-s
8	281	14.9	512	2 UC1446	serine/threonine-s
9	280	14.8	339	2 S56719	serine/threonine-s
10	277.5	14.7	502	2 T02306	probable protein k
11	277	14.6	441	2 E85362	hypothetical prote
12	277	14.6	511	1 A56009	serine/threonine-s
13	274.5	14.5	633	1 A26030	serine/threonine-s
14	274	14.5	533	1 S60304	serine/threonine-s
15	273.5	14.5	533	1 A34366	Ca2+/calmodulin-de
16	273	14.4	504	2 T07415	probable serine/th
17	272.5	14.4	480	2 A86427	probable serine/th
18	271.5	14.4	421	2 E96522	hypothetical prote
19	271	14.3	431	2 T02496	probable protein k
20	271	14.3	576	2 T41587	probable carbon ca
21	270.5	14.3	426	2 C71408	probable protein k
22	270	14.3	512	2 T07788	probable serine/th
23	269	14.2	542	1 A45025	Ca2+/calmodulin-de
24	268.5	14.2	593	1 K1ZPMN	protein kinase cdr
25	268	14.2	622	1 S44859	serine/threonine-s
26	267	14.1	472	2 B90100	SNP-related kinase
27	266	14.1	542	1 A26464	Ca2+/calmodulin-de
28	266	14.1	589	2 S68470	Ca2+/calmodulin-de
29	264	14.0	442	2 T48203	hypothetical prote

30	263	13.9	1142	2 S59359	GIN4 protein - yea
31	261.5	13.8	421	2 T48202	protein kinase AK2
32	261.5	13.8	513	1 S60303	serine/threonine-s
33	261.5	13.8	552	1 S51025	hydroxymethylglut
34	261.5	13.8	891	2 T40503	protein kinase kin
35	260.5	13.8	473	1 S59941	serine/threonine-s
36	260.5	13.8	552	1 A51621	hydroxymethylglut
37	260.5	13.8	891	2 A38903	protein kinase 1 -
38	259	13.7	350	2 T06107	probable serine/th
39	258.5	13.7	440	2 T14736	probable serine/th
40	258.5	13.7	445	2 T50802	serine/threonine-s
41	258	13.6	502	1 A41361	serine/threonine-s
42	258	13.6	726	2 T33998	hypothetical prote
43	257.5	13.6	492	1 T03321	calcium-dependent
44	257	13.6	469	2 B84644	probable protein k
45	256.5	13.6	516	1 J00270	Ca2+/calmodulin-de

ALIGNMENTS

RESULT 1

A57286 probable serine/threonine protein kinase (EC 2.7.1.-) fnk - mouse

C:Species: Mus musculus (house mouse)

C>Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 10-Sep-1997

C:Accession: A57286

R:Donohue, P.J.; Alberte, G.F.; Guo, Y.; Winkles, J.A.

J. Biol. Chem. 270, 10351-10357, 1995

A:Title: Identification by targeted differential display of an immediate early gene en

A:Reference number: A57286; PMID:95247749; PMID:7730342

A:Accession: A57286

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-631 <DON>

A:Cross-references: GB:U21392; GB:U22434

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

C:Keywords: ATP; phosphotransferase

P:61-315/Domain: protein kinase homology <KIN>

Query Match 15.8%; Score 298.5; DB 2; Length 631;
Best Local Similarity 28.3%; Pred. No. 4.5e-13;
Matches 98; Conservative 48; Mismatches 133; Indels 67; Gaps 14;

QY	39	PORLPCLPLSPPTAPRAATVATGS-RLGYYVLEBP-----	77
DB	10	PRP-PPRAVPSAPPAQPG--PPNAPSRSBEPVLAGPAPDPGRLLTDPISGRYYTK	65
QY	78	-----EGG--RAYRALHCPTEYTCVYVQESALVLEBPYALP-----	119
DB	66	GRLLGKGGFARKCYEATDTESGLIYAVKVIPOSR--VAKPHORBKILNBIELHRLQHRH	122
QY	120	VAPRTVLAQTOLLYAFTR-THGDMHSLVRSNHRIPREPAVALFROMATALAHCHQGL	178
DB	123	IYAFSHHPEDADNIYIFLELCSRKSLAHIMKANTHLLBPVRYVRLGLSKYLHQRGI	182
QY	179	VLRDLKLCRFVPRDRKKLVLENLSDCVLTGPPDSLMDKAC--PAYVGPIYSSPAS	236
DB	183	LHRDLKLGNFPIFD--NMBLKVDPGIYAARLEBPBQR--KRTTCGTFYVYAPVYTLRQG-	237
QY	237	YSGKADVMSLGVALLFTMLAGHYRFODSRPVLLFGKIRGAVYLLPGLSPARCLVCL	296
DB	238	HGPRADVMSLGVMTLLCGSPFFETADKERTYRCKQVNTYLLPSLSPARQLLAAIL	296
QY	297	RREPARLTATGTLHPMLRQ---DPMPLAPTRSHLMBAQVVPD	338
DB	297	RASPRDRPSIBQILRHDFTKGYTPDRLPVS-----SCVYVPD	334

RESULT 2
T20941
hypothetical protein F15A2.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20941
R:Gregory, J.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19349
A:Accession: T20941
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-887 <M1>
A:Cross-references: EMBL:T20207; PIDN:CA94127.1; GSPDB:GN00028; CESP:F15A2.6
A:Experimental source: clone F15A2
C:Genetics:
A:Gene: CESP:F15A.6
A:Map position: X
A:Introns: 32/1, 63/3, 92/2, 139/2, 189/3, 328/2, 448/2, 516/3, 604/2, 684/3, 735/2, 777/2

Query Match	15.8%	Score 298;	DB 2;	Length 887;
Best Local Similarity	25.9%	Pred. No. 7.1e-13;		
Matches 95, Conservative	51;	Mismatches 133;	Indels 88;	Gaps 14

Qy	62	VATSRRLRPYUV- : : :	---LEPEEGRAVYALHC : : :	PRPGTEYTCV- : : :	-----YVGBALAV : : :	107
Db	11	VAQAOYCBRYUKLEKTLGKQOTLVKGTGHCITGKAKAIKIYKKEKESBESVLQKVEBEIAI				70
Qy	108	LEPVARLEPHKIVAPTEVYLAQTOLLYAFETR-THQDMHSLVRSRRHIEPEPAVLFROM				166
Db	71	M-----KLIEHPVHLTVUYENKTKLYLLBHVSGSELFDYLVKRGRLMSKEARKPFROI				126
Qy	167	ATALAHCHONGVLVLDLTKCRFVVPADREBK- : : :	-----LVLEN- : : :	LEBSCVLTGDDSL : : :		216
Db	127	ISALDFCAHHNCCHDLKRENLLBERNNIKYADFGMASLOVEGSMLETS- : : :				177
Qy	217	WDKHCAPYVGGPEILSSRASYSGKAADVMSLVALPTMLAGHYPFODSEPFVLLFGKIRRG				276
Db	178	-----GSPHYACEVIRGE-KYDGRKADVMSCGVILYALLVGMALPFDDNNLRNLLKEVYKRG				232
Qy	277	AYALPAGLSAPARCLVRCILRRERPEARLTATGILHPWL- : : :	-----RODP- : : :	MP LAP-TRSH : : :		328
Db	223	VFHIHPYPAVDQSLIRAMIBVDPGKRYSLADVFKHPVSGTTKADPBELEPMSGVQNH				292
Qy	329	LMEAAQVY-PRDLG- : : :	-----LDBAB- : : :	EEGD : : :		351
Db	293	VIPGSDSIDPDVLRHMNCLGCFKXOKULINELLSPKRNTEKNVYFLLDRKRRAPAOBDD				352
Qy	352	REVVLYG : : :				358
Db	353	TEIVLRG 359				

RESULT 3
 S52244
 P698g3 protein - African clawed frog
 C/Specie: *Xenopus laevis* (African clawed frog)
 C/Date: 07-May-1995 #sequence_revision 03-Aug-1995 #text_change 24-May-2001
 C/Accession: S52244
 R/Roght, C.; le Guelliec, R.; Pariz, J.; Couturier, A.; Philippe, M.
 submitted to the EMBL Data Library, October 1992
 A/Description: Egg, selected by differential screening encodes a new *Xenopus* protein kin
 A/Reference number: S52243
 A/Accession: S52244
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-651 <ROG>
 A/Cross-references: EMBL:Z17205; NID:G609283; PIDN:CAA78913.1; PID:G609284
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
 C/Keywords: ATP
 P/11-265/Domain: protein kinase homology <KIN>
 P/19-27/Region: protein kinase ATP-binding motif

Query March	15.5%	Score 292.5	DB 2	Length 651
Best Local	29.9%	Pred. No. 1.2e-12		
Matches 73, Conservative	48	Mismatches 102	Indels 21	Gaps 7

```

QY      85 ALHCCTGATGTCYCKVPVQVEALATVLEPVALP-----PHKVAAPTEVLATGOLLVAF 137
Db      30 ASHLITGKVAIKIMD-KESLGDLDLPRVKTETIDAMKUNSHGVCRLYHYITFPKKI FMYL 88

QY      138 TR-THGDMSLVLRSHRIPEPEAAVLPMOMATALAHCHQGLVLRDLKLCRFVADREBK 196
Db      89 EYCPGGELTFDYIAIDRLTEEBARVFPFOIVSAVAVYIHSGQYAHNRDKKEMNTLIDEDONTL 148

QY      197 KLVLENLEDSCVLTGPDSDSLMDIKH-----ACPAVYGPRTILSRASVSGRAADVMSIGVAL 251
Db      149 KLI-----DFGLCAKPKGGL-DYHLMTCCGSAYIAAFELIOGKA-YISGBADIMSGVLT 201

QY      252 PTMLAGHYPPDQSEVFLFGKIRGAVALLPAQLSAPARCLVCLLRBPABRITATGILL 311
Db      202 YALMGYLPFPDDDNVWVLYKKIMRGKYEIIPKWLSPGSVLLSQMGQVDPDKRITVYKHLIN 261

QY      312 HPML 315
Db      262 HPML 265

```

```

RESULT 4
S72513
FOG2 protein - yeast (Kluyveromyces marxianus var. lactis)
CISpecies: Kluyveromyces marxianus var. lactis, Candida sphaerica
CIdate: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 07-May-1999
CAccession: S72513
COffprint, P.; Pizarro, A.; Domini, C.; Lodi, T.; Puglisi, P.P.; Ferrero, I.
Curr. Genet. 29, 316-326, 1996
ATitle: FOG1 and FOG2 genes, required for the transcriptional activation of glucose-t
AReference number: S72513; PMID:8598052
AAccession: S72513
AStatus: not compared with conceptual translation
AMolecule type: DNA
AResidues: 1-602 <GDP>
ANote: The source is designated as Kluyveromyces lactis
CIGenetics:
AGene: FOG2
CFunction:
AIDescription: probably involved in the regulation of glucose-repressible gene express
CISuperfamily: AMP-activated protein kinase; protein kinase homology
F13-286/Domain: protein kinase homology <KIN>

Query_Match      15.4%; Score 291; DB 2; Length 602;
Beat Local Similarity 28.5%; Pred. No. 1,4e-12;
Matches      88; Conservative 51; Mismatches 136; Indels 32; Gaps 10;

```

```

QY 63 ATARLGGVYLLER-EEG--GRAVRLALCPGTETCYKVPQOZALAVLE-----P 110
Db 27 AOGHIGKQYI KTLBEGSPGKVKLAHYHISQOYKALMIIN-KYLAASDMQGRITERIS 85
QY 111 YARLPKHQVAPRPEVYLAGTOLLAAFFRTTGDGMSLYRSRRIRTEPEAAVLPRQMATYL 170
Db 86 YLRLLRPHHIIKLDVYIKSCDEIIMVIEYAGNELEDTYVQORQMBQOEARREFQOITISAV 145
QY 171 AHCHQHGVLRLDKLCRFVPADRRERKKVLLENLEDSCVLTGPDDSLMDGAC--PAYVAP 228
Db 146 DYCGRHKIKYHRDLRPNBLLDENHLLVVKADPGL--SNMT---DGNPLKTSGSGSPRYAP 200
QY 229 EILSSRSYSGSKADWMSGLVATLTMLAGHIFPDQSEVVLFGKTRRQAYLPAQLSAPA 288
Db 201 EYISGKI-YAGPEVDWSSGVLTVMLCRRLPFDEBISPLVFKNISNGYTTIPNLSQGA 259
QY 289 RCLVRCLLRREPAERLTATGILRLPMLROD--PMLPAPRSHLMBAAQVVPDGLGDEAR 347
Db 260 ASLIKQMLIVNPPVVRITVTHIMQDQBWFRVDYLRDYLVRASHTQENS-----ESKT 309
QY 348 EEGDRREVYL 356
Db 310 EDGGPSVPL 318

```

RESULT 5


```

Db      32 GKXIAEAVVTGKHAIKILNRKIKNMEMBEKREIKIL-----RLFMPIHILROYEVI 87
      128 AGQQLYAFPTRT--HGMSLVSRRIRPEBAVLFROMATLALHCHQGLVRLDK-- 184
      88 ETTSDIYLVMEYVNSGSLFPIYVEKGLQSDERANFPQOIIISGVYCHRMNVNRDLKPE 147
      185 -----LCRFVADRRKKLVLEN--LEDSCVLTGPDSDSLMDKHAICPAYVGPETLSGRAS 236
      148 NLLDSKCNVKKIADPGLSNIMRDGHFLKTS-----GSPNYAFAVYISGKL- 193
      237 YSGKADVMSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYVALLPAGLSAPARCLVRL 296
      194 YAGPEVDVMSGCVILYVALLCGTLPPDDENIPNLFKIKGGIYTLPSHLSEARDLIPRL 253
      297 RREPARLRTATGILLHPWLRODPMPLAPTRSHLMEAAQVVPDGLDEARER--REGDREV 354
      254 IVDPMKRVITIPETIRQHPW-----QHLPRYLAVPDPPT-VQQAKKINEIYQSV 302
      355 LVYG 358
      303 VNMG 306

```

RESULT 8

```

JCI446
serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana
N/Alternate names: protein kinase SNF1 homolog
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jun-1999
C/Accession: JCI446; S58266; S66334
R/Author: L.; Thomas, M.; Blanchi, M.; Halford, N.G.; Kreis, M.
Gene 120, 249-254, 1992
A/Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein
A/Reference number: JCI446; MUID:93013041; PMID:1339373
A/Accession: S58266
A/Molecule type: DNA
A/Residues: 1-512 <LHG>
A/Cross-references: GB:M93023; NID:G166599; PIDN:AAA32736.1; PID:G166600
R/Author: F.; Kitchner, M.; Teuber, R.; Dittlich, P.
Submitted to the EMBL Data Library, May 1995
A/Description: Differential accumulation of the transcripts of 22 novel protein kinase g
A/Reference number: S58266
A/Accession: S58266
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 144-198 <THU>
A/Cross-references: EMBL:X86966; NID:G928909; PIDN:CAA60529.1; PID:G928910
R/Author: F.; Kitchner, M.; Teuber, R.; Dittlich, P.
Plant Mol. Biol. 29, 551-565, 1995
A/Title: Differential accumulation of the transcripts of 22 novel protein kinase genes 1
A/Reference number: S66314; MUID:96123233; PMID:8534852
A/Accession: S66314
A/Molecule type: DNA
A/Residues: 144-198 <TH2>
A/Cross-references: EMBL:X86966; NID:G928909; PIDN:CAA60529.1; PID:G928910
C/Comment: This enzyme plays an important role in a signal transduction cascade regulat
C/Genetics:
A/Genes: AK110; AK21
A/Introns: 64/1, 125/3, 186/3, 230/3, 292/3, 322/3, 350/3, 396/3, 475/3
C/Function: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C/Superfamily: AMP-activated protein kinase; protein kinase homolog
C/Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F/1-237/Domain: protein kinase homolog <KIN>
F/25-33/Region: protein kinase ATP-binding motif
F/146-67,147,144/Active site: Lys, Glu, Asp, Lys #status predicted
F/147,151/Binding site: magnesium (Asn, Asp) #status predicted

```

Query Match 14.9%; Score 281; DB 1; Length 512;
 Best Local Similarity 28.7%; Pred. No. 5.6e-12;
 Matches 87; Conservative 46; Mismatches 118; Indels 52; Gaps 9;

```

      80 GRAYRALHCEPTGETYCKY-----DYQALAVLEPYARLPKHKVAPTEVL 127
      31 GRVXIAEHALTGHKVAIKILNRKIKNMEMBEKREIKIL-----RLFMPIHILROYEVI 86
      128 AGQQLYAFPTRT--GDMSLVSRRIRPEBAVLFROMATLALHCHQGLVRLDK-- 184
      87 EPTDITLVMEYVNSGSLFPIYVEKGLQSDERANFPQOIIISGVYCHRMNVNRDLKPE 146
      185 -----LCRFVADRRKKLVLEN--LEDSCVLTGPDSDSLMDKHAICPAYVGPETLSGRAS 236
      147 NLLDSKCNVKKIADPGLSNIMRDGHFLKTS-----GSPNYAFAVYISGKL- 192
      237 YSGKADVMSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYVALLPAGLSAPARCLVRL 296
      193 YAGPEVDVMSGCVILYVALLCGTLPPDDENIPNLFKIKGGIYTLPSHLSEARDLIPRL 252
      297 RREPARLRTATGILLHPWLRODPMPLAPTRSHLMEAAQV--VPPDGLDEARERREGDREV 355
      253 VVDPMKRVITIPETIRQHPW-----QHLPRYLAVPDPPTVQQAKKIDEIYQSV 302
      356 LVYG 358
      303 NMG 305

```

RESULT 9

```

S56719
serine/threonine-specific protein kinase SPK-1 (EC 2.7.1.-) - soybean
C/Species: Glycine max (soybean)
C/Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 18-Jun-1999
C/Accession: S56719
R/Author: P.G.; Yoon, H.W.; Jeong, Y.H.; Bahk, J.D.; Hong, J.C.; Cho, M.J.
Submitted to the EMBL Data Library, January 1993
A/Description: Cloning of a novel protein kinase serine/threonine kinase cDNA from soybean.
A/Reference number: S56719
A/Accession: S56719
A/Molecule type: mRNA
A/Residues: 1-339 <SHI>
A/Cross-references: EMBL:L01453; NID:G169990; PIDN:AAA33979.1; PID:G169991
C/Superfamily: kinase-related transforming protein; protein kinase homolog
C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F/3-261/Domain: protein kinase homolog <KIN>
F/1-19/Region: protein kinase ATP-binding motif

```

Query Match 14.8%; Score 280; DB 2; Length 339;
 Best Local Similarity 29.3%; Pred. No. 4.2e-12;
 Matches 86; Conservative 35; Mismatches 85; Indels 88; Gaps 11;

```

      117 HKNVAPTEVLAGTQLYAFPTRT-----GDMSLVSRRIRPEBAVLFROM 166
      58 HPNIIRPEKV-----FLPTHLAIVLEVAAGSLFPRICNAGRLSDBEARFFQQL 108
      167 ATALAHCQHGVLRLDKL-----CRFPVADRRKKLVLENLSDSCVLTG 211
      109 ISGVSYCHSQICRDKLENTLLDGNPARLKCIDFGF---KSLALHSQPSVWGT- 163
      212 PDDSLMDKHAICPAYVGPETLSGRASYSKADVMSLGVALLFTMLAGHYPPQDSEPVLLF- 270
      164 -----PAYIAEVL--SRKEVGVKADVMSGCVILYVMLVGAVPFDPDPKXFR 211
      271 ----GKTRGAYALP--AGLSAPARCLVRLRREPARLRTATGILLHPWLROD--PMVL-- 322
      212 KSIGRIMSVQYALPDDYVAVSKCRHLISCFVAVPAARISIBIKOHLMPFKNLPRIIE 271
      323 -----APTRSHLMEAAQV--VPPDGLDEARER 348
      272 PBRGVDYDHKRTQPSQSVSETRMIIQEARKIKHTGSDAGTGTSDAVVAGDANER 325

```

RESULT 10
 702306
 probable protein kinase (imported) - Arabidopsis thaliana
 N/Alternate names: hypothetical protein P1317.2

C:Species: Arabidopsis thaliana (mouse-ear-cress)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C/Accession: T02306; D84753
R/Rounsbey, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, M.; Koo, H.; Moffat, K.S.; Rounsbey, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Emsw, D.; Mierman, M.C.; White, O.; Eisten, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J. Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4420; MUID:20083487; PMID:10617197
A/Accession: D84753
A/Status: preliminary
A/Molecule type: DNA
A/Accession: 1-502 <STO>
A/Cross-references: EMBL:AC004481; NID:G3337347; PIDN:AA27394.1; PID:G3337349
A/Cross-references: source: cultivar Columbia
R/Lin, X.; Kaul, S.; Rounsbey, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Emsw, D.; Mierman, M.C.; White, O.; Eisten, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J. Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4420; MUID:20083487; PMID:10617197
A/Accession: D84753
A/Status: preliminary
A/Molecule type: DNA
A/Accession: 1-502 <STO>
A/Cross-references: GB:AE002093; NID:G3337349; PIDN:AA27394.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g34180; F13P17.2
A/Map position: 2
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology <Kin>
F:155-311/Domain: protein kinase homology <Kin>

```

Query Match 14.7% Score 277.5; DB 2; Length 502;
Beet Local Similarity 26.0%; Pred. No. 9.6e-12;
Match 89; Conservative 59; Mismatches 193; Indels 61; Gaps 12;

Qy 3 ATPLAAP-----AGSLRKRLLELDNDLTERPVOKARSGPRLPCLLPUSPT 54
   ||||| : : : : : |||
Db 6 STPLAIPGPTPIPMAGLLAR--IVTKNTMETSPSPRS-----PRTP----- 48
   ||||| : : : : : |||

Qy 55 ADRATAVATASRLGIPVLLPEBEGRAYRALHCEPTGYTCQYV-----VQ 102
   ||||| : : : : : |||
Db 49 ---QGSILMDKYEIGK--LLGHGSAFYLLARNHISGSDVAIKVIDKEKIVKSLGHAIK 103
   ||||| : : : : : |||

Qy 103 EALAVLSEPVARLREPHKQVAPRTEVLAGTQLLYAPTRTH--GDMSLVNBRHRIPEEAAV 161
   ||||| : : : : : |||
Db 104 REISIL----RRVRHPRIYVHLLEVMATKTKIYIMEYVGGSELVYTV--ARGRLREGTAR 158
   ||||| : : : : : |||

Qy 162 LEROMATALAHCHQHGLVLRDLKLCRFVADBRKCLVLNLEDSCVLTGPDSLMDKXA 221
   ||||| : : : : : |||
Db 159 YEQQLISSVAFCPSRGVYHRHDKLENLLDDKGNKV-----SPGSLSVSEQLKQEGI 212
   ||||| : : : : : |||

Qy 222 C-----PAYVGEILSSRPSYSGKAADWSLGVALLFTMLAGHYRPODSEPVLLPFGKIR 275
   ||||| : : : : : |||
Db 213 CQTPCCTPAYLAVEVL--TRKGVEGAKADIMSGCIVLFTVLMAGYLFPPDDKNILVMTKIYK 271
   ||||| : : : : : |||

Qy 276 GAYALPAGLSAPRACLVRCLLRREPARRLTGTGLLHWLMO 317
   ||||| : : : : : |||
Db 272 GQPKCEKWPSPBLARLVTRMLDNTNPDRITRIPELMKHPFK 313
   ||||| : : : : : |||

```

```

RESULT 11
E85362
hypothetical protein AT4g30960 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C/Accession: E85362
R/Anonymous: The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617158
A/Accession: E85362
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-441 <STO>

```

```

A:Cross-references: GB:NC_001268; MID:g7265998; PIDN:CA879814.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4G30960
A:Map position: 4
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
Query Match 14.6%; Score 277; DB 2; Length 441;
Best Local Similarity 29.4%; Pred. No. 9e-12;
Matches 77; Conservative 39; Mismatches 124; Indels 22; Gaps 4;

QY 81 RAYRALHCPPTGYEYCKY-----PYQELALVLSPYARLRLPHKVAAPTEYLA 128
DB 37 KYTHARNIQTGKSVAMKVGKSKVGVGWDQKREISV-----RWVGHENIVELHEWMA 92
QY 129 GTQLLYAFTTRHGMNSLVRSRHRIPREBAVLFRQMATALAHCHQGLVLRDLKICRF 188
DB 93 SKSKTYFAMELVRGSELPAKAVAKGRLEBVARVYFOOLLISAVDFCHSRGTYHRDLKPENL 152
QY 189 VFADERRKLVLENLEDSCVLTGPPDSLMDKACAPVYVGEPISSASYSGGKADVWSLG 248
DB 153 LIDSEGNLKVTDGFSAPFTEHLKQDGLLTTCGTPTVYVAPEVILKK-GYGAQAKADLWSCG 211
QY 249 VALFTMLAGHYVPQDSSBPLYLFGKIRGAYVALPAGISAPARCIVRCILRRBPARTATG 308
DB 212 VILFPLLAGVLPQDDNLVNMTRKIYRGDFCKPCGWLSSDARRRLVTQLLDNPNPTIRITBK 271
QY 309 ILHFWTLRQDPW-----PLAPT 325
DB 272 VMDSFWPKQATRSRNEPVAAT 293

```

RESULT 12
A56009
serine/threonine-specific protein kinase (EC 2.7.1.-) NPK5 - common tobacco
C/Species: Nicotiana tabacum (common tobacco)
C/Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C/Accession: A56009
R/Muranaka, T.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994.
A/Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces cerevisiae
A/Reference number: A56009; MUID:94217693; PMID:8164654
A/Accession: A56009
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-511 <MUR>
A/Cross-references: GB:D26602; NID:G496384; PIDs:BA05649.1; PID:G496385
C/Function:
C/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-phosphate
C/Superfamily: AMP-activated protein kinase; protein kinase homology
C/Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F/17-27/Domain: protein kinase homology <KIN>
F/15-33/Region: protein kinase ATP-binding motif
F/48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F/147,151/Binding site: magnesium (Asn, Asp) #status predicted

	Query Match	14.6%	Score 277	DB 1	Length 511
	Similarity	28.3%	Pred. No. 1.1e-11		
	Matches	86	Conservative 49	Mismatches 115	Indels 54
					Gaps 10
Qy	80 GRAYRALHCPSTGYTCRVY-----PVOBALAVLEPYARLPKHKVAPRTBYL	127			
Db	31 GKVIASHHTLGHVAVAKILNRKKIKONMEMBEKVRREIKIL----RLPHNPHIRILEYV	86			
Qy	128 AGTOLLAAFTFRT-HGDMHSLVRSRHRIPPEBAAVLFPOMATALAHCHQGLVLRDKLC	186			
Db	87 ETPSDIVVMBYVVGSELFDYIVKEGRLOQBDEAKRFPOOISGVSEYCHRMVVRHRLDKE	146			
Qy	187 RFPVADERKKLV---LEN-----LEDSCVLVGPDDSLMDKHCAPAVGPEILSSRAS	236			
Db	147 NLLDSKMKVYKIDFGSLNTRDGHFLTSC-----GSPNPAABEVISGKL-192				
Qy	237 YSGKADVMSLGVALFTMLAGHYVPDDEPVLTFGKIRGAYVALPAGLSAPRCVLRCL	296			

```

Db      193 YAGPEVDVWSCGVILYALCGTLPDDENINPLFKCKKGGMISPSHLSAARLDLIPRML 252
OY      297 RREABRLTATGILLHPMLRQDPMLPRLATRSHLWMAQVDPDGLDEAREEBCD--REV 354
Db      253 IVDPMKRTTIPETIKMHPF-----QAHLPYLAVPPTD--MQAKKIDBDILOEV 301

OY      355 VLYG 358
Db      302 VKRG 305

RESULT 13
serine/chreonine-specific protein kinase (EC 2.7.1.-) SNF1 - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YDR477w
C/Species: Saccharomyces cerevisiae
C/Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #text_change 23-Mar-2001
C/Accession: A26030, S69644
R/Coleanza, J.L.; Carlson, M.
Science 233, 1175-1180, 1986
A/Title: A yeast gene that is essential for release from glucose repression encodes a pr
A/Reference number: A26030, MUID:86289463; PMID:3526554
A/Accession: A26030
A/Molecule type: DNA
A/Residues: 1-633 <CEU>
A/Cross-references: EMBL:M13971, NID:g172629, PIDN:AAA3058.1, PID:g172630
R/Dietrich, P. S.
submitted to the EMBL Data Library, August 1995
A/Description: The sequence of S. cerevisiae comide 9410, 8035, 8166, and 9787.
A/Reference number: S69554
A/Accession: S69644
A/Molecule type: DNA
A/Residues: 1-633 <DIR>
A/Cross-references: EMBL:U33050, NID:g927726, PIDN:AAB64904.1, PID:g927732, GSPDB:GN0000
C/Genetics:
A/Gene: SGD:SNF1, MIPS:YDR477w
A/Cross-references: SGD:S0002885, MIPS:YDR477w
A/Map position: 4R
A/Function:
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A/Note: required for expression of glucose-repressed genes in response to glucose depriv
C/Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransfe
F/51-306/Domain: protein kinase homology <KIN>
F/161-69/Region: protein kinase ATP-binding motif
F/184,103,177,179/Active site: Lys, Glu, Asp, Lys #status predicted
F/188,186/Binding site: magnesium (Asn, Asp) #status predicted
F/210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predicted

Query Match      14.5%, Score 274.5, DB 1, Length 633,
Best Local Similarity 27.8%, Pred. No. 2e-11,
Matches 88, Conservative 53, Mismatches 138, Indels 37, Gaps 11,

OY      51 SPPAPDAATVATASRGVYLLER-BEG--GRAYALHCPGTGRTYCKYVYVQGEALAV 107
Db      35 SNSTLNKSSSLADCAHGNTOYVTLGEGSGFKYKLAHYHTTGKVALKIN--KVLA 93

OY      108 LE-----PYALPPHKIVARETEVLAQTQLYAFRTTGDMSLVSRHRIPPE 158
Db      94 SDMGRIERISYLRLLHPPHILIKYDIKSDKEIIMVIEVAGNELFYIQRDMQSGE 153

OY      159 AAUVPQMAATLALACHQGLVRLDLKCRFPADREKGLVLENLEDSCVLTGPDSDAMD 218
Db      154 AARPFOQIISAVEYCHRIKHVRLDKPENLLDELINVKIDFGI--SNMT--DGNFL 208

OY      219 KGAC--PAYVGPBILSRASYSGKAADVMSLGVALLFTMLAGHYPPQDSBPVLLFGKIRG 276
Db      209 KTSGSGPPTAAPEVYSGGL-YAGPEVDVWSCGVILYALCGTLPDDENINPLFKCKKGGMISPSHLSAARLDLIPRML 267

OY      277 AYALPAGLSAPARCLVRLRREPARLTATGILLHPMLRQDPMLPRLATRSHLWMAQV 335
Db      268 VYTLPKFSLPGAAGLIKMLIVNPLNRISIHIMQDDWPKVDLPYLLP----- 316

```

```

OY      336 VPDGLGDEAREEBCD 351
Db      317 -PD---LKPRHEBENE 328

RESULT 14
serine/chreonine-specific protein kinase (EC 2.7.1.-) BKIN12 (version 2) - barley
C/Species: Hordeum vulgare (barley)
C/Date: 19-Mar-1997 #sequence_revision 15-Aug-1997 #text_change 11-Jun-1999
C/Accession: S60304, S24579
R/Halford, N.G.; Vicente-Carballo, J.; Sabelli, P.A.; Shewry, P.R.; Hannappel, U.; Kr
Plant J. 2, 791-797, 1992
A/Title: Molecular analyses of a barley multigene family homologous to the yeast prote
A/Reference number: S60303, MUID:93258420; PMID:1302632
A/Accession: S60304
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-513 <HAL>
A/Cross-references: EMBL:X65604
R/Halford, N.G.
submitted to the EMBL Data Library, April 1992
A/Reference number: S24578
A/Accession: S24579
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-61, 'A', 63-513 <HA2>
A/Cross-references: EMBL:X65604, NID:g18933, PIDN:CAA4654.1, PID:g18934
C/Function:
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C/Superfamily: AMP-activated protein kinase; protein kinase homology
C/Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kina
F/115-272/Domain: protein kinase homology <KIN>
F/23-31/Region: protein kinase ATP-binding motif
F/46,65,143,145/Active site: Lys, Glu, Asp, Lys #status predicted
F/148,152/Binding site: magnesium (Asn, Asp) #status predicted

Query Match      14.5%, Score 274, DB 1, Length 513,
Best Local Similarity 31.7%, Pred. No. 1.7e-11,
Matches 79, Conservative 42, Mismatches 102, Indels 26, Gaps 7,

OY      117 HKVAPRPVLAQTQLYAFTR--THQDMSLVSRHRIPPEBAVLPQMAATLALCHQ 175
Db      77 HPHIRYEVLETQKDFVMEYCNGBELDIYIENGLODBEARIPQOILAGVEYCHR 136

OY      176 HGLVRLDLKCRFPADREKGLV--LENLEDSCVLTGPDSDAMDKAQPA--YVGPBI 230
Db      137 IMVYHRDLKPEENLLDSRYNKLADPGLSNV-----KDGHPKTSKCSLTYAABRI 188

OY      231 LSSPASYSGKAADVMSLGVALLFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARC 290
Db      189 ISSRL-YAGPEVDVWSCGVILYALCGTLPDDENINPLFKCKKGGTILPSYLSDSARD 247

OY      291 LVRLCLRRPARLRTATGILLHPMLRQDPMLPRLATRSHLWMAQVDPDGLDEAREE 349
Db      248 LIPLDLIDPEKRIITFHIRVHPKXRLPCYLAVPPPYKQAKMI-----DSD 297

OY      350 GDRFVLYG 358
Db      298 ILRRVNLG 306

RESULT 15
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) II delta chain - rat
N/Contents: Ca2+/calmodulin-dependent protein kinase II delta chain, various splice fo
C/Species: Rattus norvegicus (Norway rat)
C/Date: 08-Jun-1990 #sequence_revision 08-Jun-1990 #text_change 11-Jun-1999
C/Accession: A34366, S39023, S39024, S39025, S39026, S39027, C47170
R/Tobimatsu, T.; Fujisawa, H.
J. Biol. Chem. 264, 17907-17912, 1989
A/Title: Tissue-specific expression of four types of rat calmodulin-dependent protein

```

Search completed: January 15, 2004, 14:56:52
Job time : 21 secs

A/Reference number: A34366; MUID:90036861; PMID:2553697
A/Accession: A34366
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-533 <TOB>
A/Cross-references: GB:J05072; NID:G203266; PIDN:AAA40866.1; PID:G203267
R/Mayer, P.; Mochly, M.; Schatz, H.; Pfeiffer, A.
PDBS Lett. 333, 315-318, 1993
A/Title: New isoforms of multifunctional calcium/calmodulin-dependent protein kinase II.
A/Reference number: S39023; MUID:94039784; PMID:8224201
A/Accession: S39023
A/Molecule type: mRNA
A/Residues: 318-328,363-371,504-533 <MAY>
A/Accession: S39024
A/Molecule type: mRNA
A/Residues: 318-371,504-511, 'N' <MA2>
A/Accession: S39025
A/Molecule type: mRNA
A/Residues: 318-328,363-371,504-511, 'N' <MA3>
A/Accession: S39026
A/Molecule type: mRNA
A/Residues: 318-349,364-371,504-533 <MA4>
A/Accession: S39027
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 318-349,364-371,504-511, 'N' <MA5>
R/Schworer, C.M.; Rothblum, L.I.; Thekkumkara, T.J.; Singer, H.A.
J. Biol. Chem. 268, 14443-14449, 1993
A/Title: Identification of novel isoforms of the delta subunit of Ca2+/calmodulin-depend
A/Reference number: A47170; MUID:93300844; PMID:8390994
A/Accession: C47170
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 314-349,364-368 <SCH>
A/Cross-references: GB:L13408; NID:G349086
A/Experimental source: skeletal muscle
A/Note: sequence extracted from NCBI backbone (NCBIN:134450, NCBIPI:134453)
C/Complex: heteromultimer composed of 10-12 alpha, beta, gamma, and delta chains general
C/Function:
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A/Note: acts on a variety of intracellular proteins; gamma and delta chains are expressed
C/Superfamily: Ca2+/calmodulin-dependent protein kinase II; protein kinase homology
C/Keywords: ATP; autophosphorylation; calmodulin binding; phosphoprotein; phosphotransfe
P.12-272/Domain: protein kinase homology <KIN>
P.20-28/Region: protein kinase ATP-binding motif
P.287-311/Region: calmodulin binding #status predicted
P.43,61,136,138/Active site: Lys, Glu, Asp, Lys #status predicted
P.1287/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status predicted

Query Match 14.5%; Score 273.5; DB 1; Length 533;
Best Local Similarity 28.8%; Pred. No. 1.9e-11;
Matches 79; Conservative 53; Mismatches 123; Indels 19; Gaps 7;

OY 61 AVATASRLGPPVLLPEBEGRAY---RALNCPGTETCYKYPVQALAL---VLEPYA 112
DB 3 STTCTRTDEQYLFEBLQKGAFSVVRRCMKIPTQGEYAKIINTKLSARDHQKLEREA 62
OY 113 ---RLRPKHVARPFEVLAAGTQLYAFTR--THGDMHSLVRSRHRIPREBAVLFRQMAT 168
DB 63 RICRLKRPNTVRLHDSISBEGFNHLPDLVTGSELPEDIVAREYSEADASHCTQQLIE 122
OY 169 ALAHCHQGLVLRDLKCRFVPADRERKLV--LENLDSCVLTGPRPDSIMDKACPAYVG 227
DB 123 SVNCHLNGIYHRDLPENLLASKSKAANVLADFGLAIVQSDQAMFGPAGTPTGIS 182
OY 228 PEILSSRLSYSGKADVWSLGVALTMTLAGHYPPDSEFVLLFGKIRGAYALPA---G 283
DB 183 PEVL--RKDPYKRPVDMACGVILYLIVGYRPFWDSDQNRLYOOIKAGAYDFSPBMDT 240
OY 284 LSPARCLVRGLRREPRARLTATGILHPRMDQ 317
DB 241 VTPEAKDLINQMLTINPAKRITASBALKHPNICO 274

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: January 15, 2004, 14:57:57 ; Search time 38 seconds
(without alignments)
1926.428 Million cell updates/sec

Title: US-09-909-474D-2

Perfect score: 1891
Sequence: 1 MRATPLAARAGSLSRKKRL.....GLGLDEAREEGDREVLYNG 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5
Searched: 762491 seqs, 20461190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/prodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/2/pubppa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/2/pubppa/US10C_PUBCOMB.pep:*
15: /cgn2_6/prodata/2/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/2/pubppa/US60_NEW_PUBCOMB.pep:*
18: /cgn2_6/prodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1887	99.8	358	US-09-799-875-8	Sequence 8, Appl1
2	1675	98.6	360	US-10-024-828-9	Sequence 9, Appl1
3	1102	58.3	233	US-09-925-101-1102	Sequence 1102, Ap
4	681.5	36.0	269	US-10-291-172-1304	Sequence 680, App
5	659.5	34.9	280	US-10-291-172-680	Sequence 680, App
6	578.5	30.6	206	US-10-228-263-2	Sequence 2, Appl1
7	408	21.6	138	US-09-864-761-45767	Sequence 45767, A
8	354	18.7	153	US-09-925-101-1367	Sequence 1367, Ap
9	304.5	16.1	373	US-10-026-021-4	Sequence 4, Appl1
10	297.5	15.7	778	US-10-354-358-92	Sequence 92, Appl1
11	297.5	15.7	778	US-10-116-325-2	Sequence 2, Appl1
12	290.5	15.4	607	US-09-769-970-15	Sequence 15, Appl1
13	290.5	15.4	607	US-10-204-041-16	Sequence 16, Appl1
14	290.5	15.4	607	US-10-108-580-2	Sequence 2, Appl1
15	287	15.2	1518	US-09-801-368-152	Sequence 152, App

16	287	15.2	1518	12	US-10-369-493-22243	Sequence 22243, A
17	283.5	15.0	651	9	US-09-870-937-10	Sequence 10, Appl1
18	283.5	15.0	651	10	US-09-974-298-112	Sequence 112, App
19	283.5	15.0	651	12	US-10-354-358-8	Sequence 8, Appl1
20	283.5	15.0	651	12	US-10-295-027-506	Sequence 506, App
21	283.5	15.0	651	12	US-10-173-999-127	Sequence 127, App
22	280	14.8	440	12	US-10-369-493-3864	Sequence 3864, Ap
23	278	14.7	290	9	US-09-867-550-1772	Sequence 1772, Ap
24	274.5	14.5	446	9	US-09-824-735-2	Sequence 2, Appl1
25	274.5	14.5	633	9	US-09-824-735-3	Sequence 3, Appl1
26	274.5	14.5	633	10	US-09-801-368-338	Sequence 338, App
27	274.5	14.5	633	12	US-10-369-493-1696	Sequence 1696, App
28	273.5	14.5	668	14	US-10-054-579-2	Sequence 2, Appl1
29	273.5	14.5	668	15	US-10-195-072-2	Sequence 2, Appl1
30	273.5	14.5	668	15	US-10-195-071-2	Sequence 2, Appl1
31	273.5	14.5	674	15	US-10-283-247-2	Sequence 2, Appl1
32	271.5	14.4	664	12	US-10-288-798-18	Sequence 18, Appl1
33	269.5	14.3	674	10	US-09-842-582-2	Sequence 2, Appl1
34	269.5	14.3	674	15	US-10-283-247-7	Sequence 7, Appl1
35	269.5	14.3	674	15	US-10-283-247-8	Sequence 8, Appl1
36	268	14.2	632	12	US-10-369-493-5806	Sequence 5806, Ap
37	266.5	14.1	499	12	US-10-354-358-88	Sequence 88, Appl1
38	266	14.1	516	12	US-09-820-790-2	Sequence 2, Appl1
39	266	14.1	542	12	US-10-820-790-4	Sequence 4, Appl1
40	265.5	14.0	703	12	US-10-116-326-6	Sequence 6, Appl1
41	263	13.9	1142	12	US-10-369-493-1700	Sequence 1700, Ap
42	261.5	13.8	436	14	US-10-006-611-2	Sequence 2, Appl1
43	261.5	13.8	891	12	US-10-369-493-2513	Sequence 2513, Ap
44	260.5	13.8	786	11	US-09-823-187-91	Sequence 91, Appl1
45	260.5	13.8	786	12	US-10-231-913-118	Sequence 118, App

ALIGNMENTS

RESULT 1
US-09-799-875-8
Sequence 8, Application US/09799875
Patient No. US20020034780A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Williamson, Mark
TITLE OF INVENTION: No. US20020034780A1el Human Protein Kinases and Uses
TITLE OF INVENTION: Theretfor
FILE REFERENCE: 35800/209996
CURRENT APPLICATION NUMBER: US/09/799,875
CURRENT FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: 60/182,059
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: 09/659,287
PRIOR FILING DATE: 2000-09-12
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 358
TYPE: PRT
ORGANISM: Homo sapiens
US-09-799-875-8
Query Match 99.8%; Score 1887; DB 9; Length 358;
Best Local Similarity 99.7%; Pred. No. 2,1e-157;
Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRATPLAARAGSLSRKKRLDNDLTPRPVQRRASGPPRLPCLLPSPPTAPRAT 60
1 MRATPLAARAGSLSRKKRLDNDLTPRPVQRRASGPPRLPCLLPSPPTAPRAT 60
QY 61 AVATASRLDPYVLLPEBEGGAYRALHCPTEYTYCKYVPVQALAVLBPYARLPKHV 120
61 AVATASRLDPYVLLPEBEGGAYRALHCPTEYTYCKYVPVQALAVLBPYARLPKHV 120
DB 61 AVATASRLDPYVLLPEBEGGAYRALHCPTEYTYCKYVPVQALAVLBPYARLPKHV 120
QY 121 ARPTVLAQTQVLLAFPTTHQDMHSLVRSRRIRPEPAVLFROMATLALHGHQGLV 180

DB 121 ARPTVLAGTQLVLAFTTRTGDMHSLVSRHRIPEPPAAVLFROMATALLAHCHQHGLVL 180
QY 181 RDLKLCRFVFPADRERKLVLENLSDSCVLTGPDSDLMDKACPAVVGPEILSSRASYSGK 240
DB 181 RDLKLCRFVFPADRERKLVLENLSDSCVLTGPDSDLMDKACPAVVGPEILSSRASYSGK 240
QY 241 AADVWSLGVALLFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRRBP 300
DB 241 AADVWSLGVALLFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRRBP 300
QY 301 AERLTATGILLHPMLRQDPMPLAFTSRHLMBAQVVPDGLGIDBARSEEGDREVLYG 358
DB 301 AERLTATGILLHPMLRQDPMPLAFTSRHLMBAQVVPDGLGIDBARSEEGDREVLYG 358

RESULT 2
US-10-024-828-9
Sequence 9, Application US/10024828
Publication No. US20030036051A1
GENERAL INFORMATION:
APPLICANT: Virca, Duke
APPLICANT: Bird, Timothy A.
APPLICANT: Anderson, Dirk M.
APPLICANT: Marken, John S.
TITLE OF INVENTION: Human CDNA5 Encoding Polypeptides Having Kinase
FILE REFERENCE: 2877-US
CURRENT APPLICATION NUMBER: US/10/024, 828
CURRENT FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: US/09/509, 902A
PRIOR FILING DATE: 1999-08-03
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
US-10-024-828-9

Query Match 88.6%; Score 1675; DB 15; Length 360;
Best Local Similarity 99.7%; Pred. No. 8.8e-139;
Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRATPLAAPAGSLSRKKRLLELDNLDTERRPVQKARSGSPQRLRPCLPLSPPTAPDRAT 60
DB 43 MRATPLAAPAGSLSRKKRLLELDNLDTERRPVQKARSGSPQRLRPCLPLSPPTAPDRAT 102
QY 61 AVATASRLGPPVLLPEPEGRAVYALHCPGTGTEYTCVVYQGEALAVLEPYARLPKHV 120
DB 103 AVATASRLGPPVLLPEPEGRAVYALHCPGTGTEYTCVVYQGEALAVLEPYARLPKHV 162
QY 121 ARPTVLAGTQLVLAFTTRTGDMHSLVSRHRIPEPPAAVLFROMATALLAHCHQHGLVL 180
DB 163 ARPTVLAGTQLVLAFTTRTGDMHSLVSRHRIPEPPAAVLFROMATALLAHCHQHGLVL 222
QY 181 RDLKLCRFVFPADRERKLVLENLSDSCVLTGPDSDLMDKACPAVVGPEILSSRASYSGK 240
DB 223 RDLKLCRFVFPADRERKLVLENLSDSCVLTGPDSDLMDKACPAVVGPEILSSRASYSGK 282
QY 241 AADVWSLGVALLFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRRBP 300
DB 283 AADVWSLGVALLFTMLAGHYPPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRRBP 342
QY 301 AERLTATGILLHPMLRQD 318
DB 343 AERLTATGILLHPMLRQD 360

RESULT 3
US-09-925-301-1102
Sequence 1102, Application US/09925301
Patent No. US20020052308A1

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925, 301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124, 270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1102
LENGTH: 233
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-301-1102

Query Match 58.3%; Score 1102; DB 9; Length 233;
Best Local Similarity 100.0%; Pred. No. 1e-88;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 151 RHRIPPEBAVLFROMATALLAHCHQHGLVLRDLKLCRFVFPADRERKLVLENLSDSCVLT 210
DB 26 RHRIPPEBAVLFROMATALLAHCHQHGLVLRDLKLCRFVFPADRERKLVLENLSDSCVLT 85
QY 211 GPDSDLMDKACPAVVGPEILSSRASYSGKADVWSLGVALLFTMLAGHYPPQDSBPVLLF 270
DB 86 GPDSDLMDKACPAVVGPEILSSRASYSGKADVWSLGVALLFTMLAGHYPPQDSBPVLLF 145
QY 271 GKIRGAYALPAGLSAPARCLVRCILRRPAPERTATGILLHPMLRQDPMPLAFTSRHLM 330
DB 146 GKIRGAYALPAGLSAPARCLVRCILRRPAPERTATGILLHPMLRQDPMPLAFTSRHLM 205
QY 331 EAAQVVPDGLGIDBARSEEGDREVLYG 358
DB 206 EAAQVVPDGLGIDBARSEEGDREVLYG 233

RESULT 4
US-10-291-172-304
Sequence 304, Application US/10291172
Publication No. US20030228584A1
GENERAL INFORMATION:
APPLICANT: Hyeq, Inc
TITLE OF INVENTION: No. US20030228584A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-045
CURRENT APPLICATION NUMBER: US/10/291, 172
CURRENT FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/693, 267
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/665, 363
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/616, 847
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 09/596, 193
PRIOR FILING DATE: 2000-06-17
PRIOR APPLICATION NUMBER: 09/574, 454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519, 705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 752
SEQ ID NO 304
LENGTH: 269
TYPE: PRT
ORGANISM: Homo sapiens
US-10-291-172-304

Query Match 36.0%; Score 681.5; DB 12; Length 269;
Best Local Similarity 51.8%; Pred. No. 1.1e-51;
Matches 133; Conservative 38; Mismatches 79; Indels 7; Gaps 2;
QY 88 CP-----TGTEYTCVVYQGEALAVLEPYARLPKHVAVPTREVLATGQLVLAFTTRTH 141

```
Db      2  CPGKCASTLGRVAVCKVPIPIQDRIKIRPIYIGLPSHSNITGIVETGETKAVYFFBKDP 61
      142  GDMHSLVRSRRIRIPEPEAAVLFROMATLALHCHQGLVLDLKCIRFPADRRKQVLE 201
      62  GDMHSLVRSRRIRIPEPEAAVLFQIVSAVAHQSAIVLGDLCIRKRFVFTSEBRTOLRL 121
      202  NLBESCVLTGDDSLMDKACPAVYVGPETILSSBASYSKADVWSIGVALFTMLAGHYPP 261
      122  SLEDTIMKGBDDALSDHGCPAVYSEILNTTGTISGKAADVWSIGVMLTYTLVGRIYP 181
      262  QDSEPVLLFGKIRGAYALPAGLSAPARCLVCLRRPABRLTATGILLHPMLRQDPM 321
      182  HDSPSALFSGKIRRGFCIPENHISPKARCLIRSLRRPESRLTAPBILHPFESVLEP 241
      322  LAPRSHLWMAQVVPD 338
      242  -GYIDSEIGTSDQIVPE 257
```

RESULT 5

```
US-10-291-172-680
/ Sequence 680, Application US/10291172
/ Publication No. US20030228584A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyeeg, Inc
/ TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
/ FILE REFERENCE: 21272-045
/ CURRENT APPLICATION NUMBER: US/10/291.172
/ PRIOR FILING DATE: 2000-11-08
/ PRIOR APPLICATION NUMBER: 09/693,267
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 09/665,363
/ PRIOR FILING DATE: 2000-09-19
/ PRIOR APPLICATION NUMBER: 09/616,847
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 09/596,193
/ PRIOR FILING DATE: 2000-06-17
/ PRIOR APPLICATION NUMBER: 09/574,454
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: 09/519,705
/ PRIOR FILING DATE: 2000-03-07
/ NUMBER OF SEQ ID NOS: 752
/ SEQ ID NO 680
/ LENGTH: 290
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-291-172-680
```

Query Match 34.9%; Score 659.5; DB 12; Length 290;

Best Local Similarity 51.2%; Pred. No. 1e-49; Mismatches 78; Indels 1; Gaps 1;

```
Matches 127; Conservative 42; Mismatches 78; Indels 1; Gaps 1;
      91  GTEYTCNVYVQEAALVLEPYARLPKHVARPTETVLATGQLVAFPTKTHGDMHSLVRS 150
      22  GPPEFSLQVPIKHYQDRIKIRPIYIGLPSHSNITGIVETGETKAVYFFBKDPGDMHSLVRS 81
      151  RHRIPEBAVLFROMATLALHCHQGLVLDLKCIRFPADRRKQVLENTDSCVLT 210
      82  RKLRREBAARLKQIVSAVAHQSAIVLGDLCIRKRFVFTSEBRTOLRLSELDTHIMK 141
      211  GPDDSLMDKACPAVYVGPETILSSBASYSKADVWSIGVALFTMLAGHYPPDSEPVLLF 270
      142  GEDDALSDHGCPAVYSEILNTTGTISGKAADVWSIGVMLTYTLVGRIYFFHSDPSALF 201
      271  GKIRGAYALPAGLSAPARCLVCLRRPABRLTATGILLHPMLRQDPMPLAFTSRSLW 330
      202  SKIRRGFCIPENHISPKARCLIRSLRRPESRLTAPBILHPFESVLEP-GYIDSEIG 260
      331  EAAQVVPD 338
      261  TSDQIVPE 268
```

RESULT 6

```
US-10-228-263-2
/ Sequence 2, Application US/10228263
/ Publication No. US2003009985A1
/ GENERAL INFORMATION:
/ APPLICANT: Li, Jing
/ TITLE OF INVENTION: AMPLIFIED GENE INVOLVED IN CANCER
/ FILE REFERENCE: 38002-0034
/ CURRENT APPLICATION NUMBER: US/10/228.263
/ PRIOR FILING DATE: 2002-12-04
/ PRIOR APPLICATION NUMBER: US 60/330,797
/ PRIOR FILING DATE: 2001-10-31
/ PRIOR APPLICATION NUMBER: US 60/314,655
/ PRIOR FILING DATE: 2001-08-27
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 206
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-228-263-2
```

Query Match 30.6%; Score 578.5; DB 15; Length 206;

Best Local Similarity 57.9%; Pred. No. 8.3e-43; Mismatches 52; Indels 1; Gaps 1;

```
Matches 113; Conservative 29; Mismatches 52; Indels 1; Gaps 1;
      144  MHSVRSRRIRIPEPEAAVLFROMATLALHCHQGLVLDLKCIRFPADRRKQVLENTL 203
      1  MHSVRSRRIRIPEPEAAVLFQIVSAVAHQSAIVLGDLCIRKRFVFTSEBRTOLRLSL 60
      204  EDCSVLTGDDSLMDKACPAVYVGPETILSSBASYSKADVWSIGVALFTMLAGHYPPD 263
      61  EDTIMKGBDDALSDHGCPAVYSEILNTTGTISGKAADVWSIGVMLTYTLVGRIYFFH 120
      264  SEPVLLFGKIRGAYALPAGLSAPARCLVCLRRPABRLTATGILLHPMLRQDPMPLA 323
      121  SDPSALFSGKIRRGFCIPENHISPKARCLIRSLRRPESRLTAPBILHPFESVLEP-G 179
      324  PTRSHLWMAQVVPD 338
      180  YIDSEIGTSDQIVPE 194
```

RESULT 7

```
US-09-864-761-45767
/ Sequence 45767, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wenheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aeomic8-X-1
/ CURRENT APPLICATION NUMBER: US/09/864.761
/ PRIOR FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
```

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 45767
LENGTH: 138
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009486.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
OTHER INFORMATION: SWISSPROT HIT: 074536, EVALU = 1.00e-18
OTHER INFORMATION: EST_HUMAN HIT: B897149.1, EVALU = 2.00e-71
US-09-864-761-45767

Query Match 21.6%; Score 408; DB 9; Length 138;
Best Local Similarity 63.1%; Pred. No. 4,6e-28;
Matches 77; Conservative 15; Mismatches 30; Indels 0; Gaps 0;

QY 197 KLVENLEDSCTLPDDSLMDKACPAVYVGEILLSSASVSGKADWWSLGVAFPTKLA 256
DB 2 RVKESLSDAYILRDDBSLSDKHCPCAVSPEILLTSGSYSGKADWWSLGVAFPTKLA 61
QY 257 GHVPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRRPAPERTLATGILLHPMLR 316
DB 62 GRYPHDIEPSLPSKIRGCPNIFETLSPKAKCLIRSLRREPSERLTSQRIIDHPMFS 121
QY 317 QD 318
DB 122 TD 123

RESULT 8
US-09-925-301-1367
Sequence 1367, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1367

LENGTH: 153
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (136)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (138)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (141)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (142)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (143)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (152)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1367

Query Match 18.7%; Score 354; DB 9; Length 153;
Best Local Similarity 93.3%; Pred. No. 2.9e-23;
Matches 70; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRATPLAPAGSLRKRRLBIDNLDTERPVQKARSQPQRLPCLPLSPPTAPDRAT 60
DB 74 MRATPLAPAGSLRKRRLBIDNLDTERPVQKARSQPQRLPCLPLSPPTAPDRAT 133
QY 61 AVATASRLGPVYLE 75
DB 134 AVXTSRXXVYLE 148

RESULT 9
US-10-026-021-4
Sequence 4, Application US/10026021
Publication No. US20030027756A1
GENERAL INFORMATION:
APPLICANT: Hitoshi Yasumichi
APPLICANT: Demo, Susan
APPLICANT: Deming, Yonchu
APPLICANT: Rigel Pharmaceuticals, Inc.
TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for
FILE REFERENCE: 021044-001210US
CURRENT APPLICATION NUMBER: US/10/026,021
CURRENT FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/309,632
PRIOR FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (1)-(373)
OTHER INFORMATION: human FNK mitotic kinase domain
US-10-026-021-4

Query Match 16.1%; Score 304.5; DB 15; Length 373;
Best Local Similarity 28.9%; Pred. No. 2.1e-18;
Matches 101; Conservative 48; Mismatches 134; Indels 67; Gaps 15;

QY 29 RPYQKAR-----SGPQRLPCL-----LPLSP-----TAPDRATAVATSRRL 68
DB 11 RPYQKARPAAGPAPGPP--PPSALRGPELEMLAGLPTSDGRITIDPDSGRITVLRGRL 68


```

OY 69 GRVYLLERBEGG--RAYRNLHCSTGIBVYCKVYRVOBALAVEPYARLP-----115
Db 69 G-----KGGPARCTEATDTGETSGAYVAKVYRGR---VVFHQREKLINEIELHRDL 117
OY 116 PKHVAPRPTVLTAQOLLYAFPTR--THGDMHSIVRSRHRIRPEBAALVFRQMATALACH 174
Db 118 QHRHIVRFSHNEDEDANITYIPLFELCSRKSLAHIMWKAHRTLLEBRYVYUYLEQIISGLKYTL 177
OY 175 OHGLVLRDLCKLRFVAPADREKKVLVLENLEDSGVLGPDDSLMDKAC--PAVVGPEILS 232
Db 178 OHGILHRDLKLGNSFFTT--ENNELKAKGDPGLAARLEPPEOR--KXII CGIPNYVABEVL 233
OY 233 SRASVSGKADVMSLGVALFTMLAGHYFPDSEBVLVFGKIRGAVALPAGLSAPARCT 292
Db 234 RQG--HGPEADVMSLGCWYITLLCGSPFPETADLKEVYRCIKQVHTLTPASLSIPARQL 291
OY 293 RCLLRBPRLRTATQIILHPMLRQ---DPMGLATPRTSHMLAAOVPRD 338
Db 292 AAILPASPDRPSIDILRHDPFKYTPRPLIS-----SCATVP 333

```

Query Match	297.5;	DB 12,	Length 778;
Best Local Similarity	27.4%;	Pred. No. 2,2e-17;	
Matches	98;	Conservative 54;	Mismatches 14;
		Indels 63;	Gaps 14

Qy	44	PPCLILSPPAAPDRATAVATASRLGPVTL---	LEPSEGGAAVRAALACPRTGTYTCR--	98
Db	13	PAYHLPHHPPHPOHAQVY-----	GPRLKRTKLGSGTGVLGVNHITQKVAIKIVN	66
Qy	99	-----YVQSLATLVEPVARLPKHVARPTEVAGTQLTYAFETR--	THGDMHSLV	148
Db	67	REKLSBVLTKVBERIALT-----	KLIEHPVTLKLDVYENKKYTLVLBNHSGGELFOYL	122
Qy	149	RSRHHIPPEPPAAVLPROMATATLACHQGVLPRLQICRFPADBERKCVLLENIDSCV	208	
Db	123	VKKGHLTPKEKAPFRQIVSADDFCHSISICHRDIAKPNLIL--	DEKNINRIADGMAASL	180
Qy	209	LTGPDLSIMDKAAC--PAYVGEPIILSRASVSGKADWVSLGVALLFTYLACHYPPQDSBP	266	

```

RESULT 10
US-10-354-358-92
? Sequence 92, Application US/10354358
? Publication No. US20030157082A1
? GENERAL INFORMATION:
? APPLICANT: Millennium Pharmaceuticals, Inc
? APPLICANT: Hunter, John Joseph
? APPLICANT: Macbeth, Kyle J.
? APPLICANT: Teal, Fong-ying
? APPLICANT: Lesoon, Andrea
? APPLICANT: Lightcap, Eric S.
? APPLICANT: Williamson, Mark
? APPLICANT: Rudolph-Owen, Laura A.
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
? TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
? TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
? TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
? TITLE OF INVENTION: 9552, 9389, 1642, 85262, 10297, 1584, 9525, 14124, 4469,
? TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
? TITLE OF INVENTION: 15334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
? TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
? FILE REFERENCE: MP102-020PIRNONMIN
? CURRENT APPLICATION NUMBER: US/10/354,358
? CURRENT FILING DATE: 2003-01-30
? PRIOR APPLICATION NUMBER: US 60/353,600
? PRIOR FILING DATE: 2002-01-31
? PRIOR APPLICATION NUMBER: US 60/364,517
? PRIOR FILING DATE: 2002-03-15
? PRIOR APPLICATION NUMBER: US 60/371,075
? PRIOR FILING DATE: 2002-04-09
? PRIOR APPLICATION NUMBER: US 60/371,507
? PRIOR FILING DATE: 2002-04-10
? PRIOR APPLICATION NUMBER: US 60/372,984
? PRIOR FILING DATE: 2002-04-16
? PRIOR APPLICATION NUMBER: US 60/374,194
? PRIOR FILING DATE: 2002-04-19
? PRIOR APPLICATION NUMBER: US 60/382,995
? PRIOR FILING DATE: 2002-05-24
? PRIOR APPLICATION NUMBER: US 60/385,023
? PRIOR FILING DATE: 2002-05-31
? PRIOR APPLICATION NUMBER: US 60/388,853
? PRIOR FILING DATE: 2002-06-14
? PRIOR APPLICATION NUMBER: US 60/389,395
? PRIOR FILING DATE: 2002-06-17
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 122
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 92
? LENGTH: 778
? TYPE: prt
? ORGANISM: Homo sapiens
? US-10-354-358-92

```

```

RESULT 11
US-10-116-326-2
Sequence 2, Application US/10116326
Publication No. US2003016689A1
GENERAL INFORMATION:
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Mathur, Brian
APPLICANT: Fritschie, Carl Johan
FILE OF INVENTION: NO. US2003016689A1el Human Kinases and Polynucleotides Encodind
FILE REFERENCE: LEX-0332-USA
CURRENT APPLICATION NUMBER: US/10/116,326
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/282,036
PRIOR FILING DATE: 2001-04-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 778
TYPE: PRT
ORGANISM: homo sapiens
US-10-116-326-2

Query Match      15.7%; Score 297.5; DB 12; Length 778;
Best Local Similarity 27.4%; Pred. No. 2,2e-17;
Matches 98; Conservative 54; Mismatches 143; Indels 63; Gaps 14;

44 PCCLLPSEPPTAPDQATAVATASRLGPYTL---LPBBSGGRAVRAHLCPTGETYTCV-- 98
13 PAYHLPHPHPHPOHAQVY-----GPYRLKLTGKGQGLVTLGVHCICNGQVAIKIVN 66
99 -----YPOBALAVLEPYARLPHNGHVARPTEVLNAGTOLVYAFPR--THGDMSLV 148
67 REKLSBSVLMKVERITAIL-----KLIRPHVLTADYENKTKLYLVLRHVSQGSLEPDYL 122
149 RSRHRIPEPBAVLFQWATALAHCHQGLVLRDLTLCFVPVADRERKLVLENTLBDSCV 208
123 VKQGLTLPEKARKPFQIYVSLDFCHSYGICHRDLCPBULL--DEKNIRIDFQWASL 180
209 LTGPDDSLMDKAA--PAYGPEILSSRASYSYGKADWVSLGYALTMLAGHYPPQDSEP 266
181 QVG--DSLLET--SCGSPHYACEVVIKG--KYDGRARDWMSCGVILTALLVGLALPFDDNTL 236
267 VLEPKIRGAVVALPAGLSAPARCLVRCILREBPARTLATGTILHPML---RQDPMPFL- 322

```

Db 237 RQLLEKVGKGVHMFIPDQOSILRGWIEVEPKRLSLBQIOKRPVYLGGKHEPDCI 296
Qy 323 --APRSHLMEAAQVPPDGLD-----BARBEGDREVLY 357
Db 297 BPAPGR--VAKRSLPENGELDPVLESMSLGGPRDRERLRLRLESEERQEMIT 351

RESULT 12
US-09-769-970-15
Sequence 15, Application US/09769970
Publication No. US20030170219A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Hillman, Jennifer L.
Corley, Neil C.
Quejler, Karl G.
Lal, Preeti
Goli, Surya K.
Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PasteSO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,970
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/272,796
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billing, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 607 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1827450
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-769-970-15

Query Match 15.4%; Score 290.5; DB 12; Length 607;
Beet Local Similarity 29.2%; Pred. No. 6.7e-17;
Matches 92; Conservative 45; Mismatches 129; Indels 49; Gaps 12;

Qy 48 LPLSP-----TAPDATAVATASRLGPRVYLLERPEG--RAYRALHCRPTGYTCKYRV 101
Db 5 LPTSDPGRLLTDPKSGRTYLLKGRLLG-----KGGPARCYEATDTETGSAYAVKVIPO 56
Qy 102 QBALAVLEPVYARLP-----PHKIVARPTVLAGTOLLVAFPTT-THGDHMSLVR 149
Db 57 SR---VAKPHQBEKILNBIELHRLDQHRHIVAFSHHFPEDADNIVYIFLELCSKSLAHITWK 113

Qy 150 SRHRIPEBEAAVLFQCATALAHCHQGLVLRDLKCFVPADEBRKKVLLENLSDCVL 209
Db 114 ARHTLLEBEVYVYLRQILSGIKYLRHGDILKDFIT--ENMELKYDDPGLAARL 171
Qy 210 TGPDDSLMDKAC--PAYGPEILSSRASYSCKAADVMSLGVALLFTMLAGHYPPQDSBPV 267
Db 172 BPBGR--KKTICGTPNVVAVPEVLRQ--HGPEADVMSLGVMTLLCGSPPEFTADLK 227
Qy 268 LPLFGIRGAYALPAGSAPARCLVRCLREPARLRTATGILLHPMLRQ----DPMPLA 323
Db 228 ETVKCIQVHTYTLPLASLSLPRQLLAAILRASPRDRPSIDQILHRDFTTKGYTPRLRPS 287
Qy 324 PTRSHLMEAAQVDP 338
Db 288 -----SCVTPD 294

RESULT 13
US-10-204-041-16
Sequence 16, Application US/10204041
Publication No. US20030176443A1
GENERAL INFORMATION:
APPLICANT: STEIN-GERLACH, MATTHIAS
APPLICANT: SALASIDIS, KONSTADINOS
APPLICANT: BACHER, GERALD
APPLICANT: MULLER, STEFAN
TITLE OF INVENTION: Pyridylpyrimidine Derivatives as Effective Compounds Against Pr
FILE REFERENCE: AWM-007.1P US
CURRENT APPLICATION NUMBER: US/10/204,041
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: EP 01111858.5
PRIOR FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: PCT/EP02/05420
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 607
TYPE: PRT
ORGANISM: Homo sapiens
US-10-204-041-16

Query Match 15.4%; Score 290.5; DB 12; Length 607;
Beet Local Similarity 29.2%; Pred. No. 6.7e-17;
Matches 92; Conservative 45; Mismatches 129; Indels 49; Gaps 12;

Qy 48 LPLSP-----TAPDATAVATASRLGPRVYLLERPEG--RAYRALHCRPTGYTCKYRV 101
Db 5 LPTSDPGRLLTDPKSGRTYLLKGRLLG-----KGGPARCYEATDTETGSAYAVKVIPO 56
Qy 102 QBALAVLEPVYARLP-----PHKIVARPTVLAGTOLLVAFPTT-THGDHMSLVR 149
Db 57 SR---VAKPHQBEKILNBIELHRLDQHRHIVAFSHHFPEDADNIVYIFLELCSKSLAHITWK 113
Qy 150 SRHRIPEBEAAVLFQCATALAHCHQGLVLRDLKCFVPADEBRKKVLLENLSDCVL 209
Db 114 ARHTLLEBEVYVYLRQILSGIKYLRHGDILKDFIT--ENMELKYDDPGLAARL 171
Qy 210 TGPDDSLMDKAC--PAYGPEILSSRASYSCKAADVMSLGVALLFTMLAGHYPPQDSBPV 267
Db 172 BPBGR--KKTICGTPNVVAVPEVLRQ--HGPEADVMSLGVMTLLCGSPPEFTADLK 227
Qy 268 LPLFGIRGAYALPAGSAPARCLVRCLREPARLRTATGILLHPMLRQ----DPMPLA 323
Db 228 ETVKCIQVHTYTLPLASLSLPRQLLAAILRASPRDRPSIDQILHRDFTTKGYTPRLRPS 287
Qy 324 PTRSHLMEAAQVDP 338
Db 288 -----SCVTPD 294

```
RESULT 14
US-10-108-580-2
; Sequence 2, Application US/10108580
; Publication No. US2003007681A1
; GENERAL INFORMATION:
; APPLICANT: Cogswell, John
; TITLE OF INVENTION: PLK3 PROTEIN-PROTEIN INTERACTIONS
; FILE REFERENCE: P04458
; CURRENT APPLICATION NUMBER: US/10/108,580
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 607
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-108-580-2

Query Match      15.4%; Score 290.5; DB 15; Length 607;
Best Local Similarity 29.2%; Pred. No. 6.7e-17;
Matches 92; Conservative 45; Mismatches 129; Indels 49; Gaps 12;

Cy      48 LPLSP-----TAPRATAVATASRLGPyVLEPERGG--RAYRALHCPGTGYTCKYVPV 101
Db      5 LPTSDPRLITDPRSGRYLKGRLG-----KGFARCYEATDTEGSAVAVKVIHQ 56

Cy      102 QEALAVLEPYARLP-----PHKVAPTEVLAGTQLLYAFETR--THGDMHSLVR 149
Db      57 SR---VAKPHQREKILNEIHLRDLOHRIIVFSGHPEADADNIYFLELCSRKSLAHIMK 113

Cy      150 SRHRIPEPEAVLFROMATALAHCHQGLVLRDLKCFVADREKKLVLENLEDSCVL 209
Db      114 ARHTLLEEVRYVYRQILSGLKYLHQGILHRDLGNFPTT--ENMELKVGDGLARL 171

Cy      210 TGPDDSLMDKAC--PAYGPEILSSRASYSKADVSLVALFTMLAGHYPPDSEPV 267
Db      172 EPPQR--KKTICGPNVAPEVLRQG--HGPAADVSLGCVWTTLCGSPFETDLK 227

Cy      268 LLEFKIRGAYALPAGISAPARCVCRLRREPRRLTATGILLHPWLQ---DPMPLA 323
Db      228 ETVRCIKOVHYTLPLASISLPLARQLLAAILRASPRDRPSIDQILRHDFPTKGYTDRLPIS 287

Cy      324 PTRSHLWBAQVDP 338
Db      288 -----SCVTFPD 294

RESULT 15
US-09-801-368-152
; Sequence 152, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Bueby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Maxon, Mary
; APPLICANT: Madden, Kevin
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amit
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
```

```
NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 152
; LENGTH: 1518
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-152

Query Match      15.2%; Score 287; DB 10; Length 1518;
Best Local Similarity 26.7%; Pred. No. 4.5e-16;
Matches 86; Conservative 50; Mismatches 128; Indels 58; Gaps 9;

Cy      56 PDRAVAVTAS-----RLGPVYL---LEPEBGSAARALHCPGTGYTCKYVPVQGL 105
Db      59 PDSTVSATKSSKKRSRDTVGPMKLGKTLGKSSGKRVRLAKMETGQLAAIKIVPKKAF 118

Cy      106 -----AVLEPYA-----RLPPKHVAR 122
Db      119 VHCNNGIVPNSSSVSWTSNVSSPSIASREHNSQTNPYGIRESIVIMKLISHTNWA 178

Cy      123 PTEVLACTQLLYAFETRTH--GDMHSLVRSRHRIPPEPAVTLFROMATALAHCHQGLVLR 181
Db      179 LFEVWENKSELVYLEYVDGGELEFDYLVSKGLPEREAIHYFKQIVEGVSYCHSFNI 238

Cy      182 DLKICRFVADREKKLVLENLEDSCVLTGPDDSLMDGACPAVYGPEILSSRASYSKA 241
Db      239 DLK--PENLLDKKRRRIKIADF--GMAALELPNKLKLTSCGSPHYASPIVMGR--DYHGCP 295

Cy      242 ADVMSLGVALLFTMLAGHYPPDSEPVLLFGKIRGAYALPAGISAPARCVCRLRREPA 301
Db      296 SDVMSGIVLPALLTGHLPFMDNIIKKLLLVQSGKYMPSNLSBADLISKIIVIDE 355

Cy      302 ERLTATGILLHPWLQ---DPMPL 322
Db      356 KRITTOBILKRPILIKYDDELVP 377
```

Search completed: January 15, 2004, 15:03:53
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 15, 2004, 14:55:46 ; Search time 44 Seconds
(without alignments)
1291.457 Million cell updates/sec

Title: US-09-909-474d-2

Perfect score: 1891
Sequence: 1 MRAIPLAPAGSLSRKKRL.....GLGIDBARBEGDREVLVYG 358

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107663 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107663

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1887	99.8	358	AAU03509	Human protein kinase
2	1884	99.6	358	AAK8908	Human polypeptide
3	1884	99.6	358	AAK20326	Human protein phosphatase
4	1884	99.6	358	AAK0694	Human polypeptide
5	1870	98.9	360	ABR0976	Human protein kinase
6	1675	88.6	360	AAK6917	Human protein kinase
7	1582.5	83.7	323	AAK6917	Human protein kinase
8	1102	58.3	233	AAK3657	Human cancer associated protein
9	1097	58.0	278	ABR06093	Human NS protein

10	1097	58.0	278	23	ABR06108	Human NS protein
11	744.5	39.4	372	22	ABR08975	Human protein kinase
12	681.5	36.0	269	22	AAU28135	Novel human secret
13	659.5	34.9	290	22	AAU28323	Novel human secret
14	578.5	30.6	206	24	ABP6856	Human C8FW protein
15	435	23.0	484	22	ABR71379	Drosophila melanog
16	408	21.6	138	22	ABG58479	Human liver peptid
17	408	21.6	138	22	ABR43079	Peptide #10585 enc
18	408	21.6	138	22	AAK63978	Human brain expres
19	408	21.6	138	22	AAK6904	Human bone marrow
20	408	21.6	138	22	ABG45957	Peptide #10941 enc
21	354	18.7	153	21	AAK39922	Human peptide enco
22	316.5	16.7	371	21	AAK54410	Human kinase
23	297.5	15.7	778	24	ABR8743	Human kinase #1.
24	297.5	15.7	778	24	ABR8743	Zea mays protein
25	297	15.7	327	21	AAK54419	Human protein kinase
26	295	15.6	794	22	AAU03517	Human kinase PKIN
27	295	15.6	794	23	ABG16271	Novel human diagno
28	291.5	15.4	605	22	ABG16826	wheat putative car
29	291	15.4	523	21	AAK03425	Human kinase
30	291	15.4	523	24	ABR40719	Human kinase
31	290.5	15.4	607	24	AAK34495	Human PRK protein
32	287.5	15.2	754	24	AAK16604	Human cell cycle-r
33	287.5	15.2	1078	24	ABP6069	Human protein kinase
34	287	15.2	512	23	AAK50578	Arabidopsis SNF-1
35	284	15.0	504	24	ABR0815	Cucumis sativus ol
36	284	15.0	512	21	AAK36157	Arabidopsis thalia
37	283.5	15.0	619	20	ABR04768	Human dominant neg
38	283.5	15.0	643	20	ABR04769	Mouse dominant neg
39	283.5	15.0	651	23	AAK47857	KIAA0175 protein
40	283.5	15.0	651	24	ABR57635	Differentially exp
41	283.5	15.0	651	24	ABR57635	Lung cancer-associ
42	283	15.0	438	21	AAK3421	Soybean putative c
43	283	15.0	438	24	ABR40715	Glycine max oil tr
44	283	15.0	512	23	AAK17663	A thaliana AKIN11
45	283	15.0	512	23	AAK17664	A thaliana AKIN11

ALIGNMENTS

RESULT 1	AAU03509	standard; Protein: 358 AA.
XX	AAU03509;	
AC	12-SEP-2001	(first entry)
DT	XX	
DB	XX	Human protein kinase #9.
XX	XX	
KW	XX	Human, protein kinase; PTK, STK; cancer; cardiovascular disease;
KW	XX	metabolic disorder; immune related disease; neurological disorder;
KW	XX	neurodegenerative disorder; inflammatory disorder; infectious disease;
KW	XX	reproductive disorder.
OS	XX	Homo sapiens.
PN	XX	MO200138503-A2.
PD	XX	31-MAY-2001.
XX	XX	
XX	XX	22-NOV-2000; 2000MO-US32085.
XX	XX	
PR	XX	24-NOV-1999; 99US-0167482.
XX	XX	(SUGR-) SUGEN INC.
PA	XX	
XX	XX	Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX	XX	Plowman P, Clary D;
XX	XX	WPI; 2001-343950/36.
DR	XX	N-PSDB; AAS06709.

XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
XX Claim 7, Figure 2, 43pp; English.
XX
XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancer (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.
XX
XX Sequence 358 AA:
SQ
Query Match 99.8%; Score 1887; DB 22; Length 358;
Best Local Similarity 99.7%; Pred. No. 26-171;
Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 MRATPLAAPAGSLSKKRLLELDNDLTERPVOKRARSQPRLPCLLPSPPTAPDRT 60
DB 1 MRATPLAAPAGSLSKKRLLELDNDLTERPVOKRARSQPRLPCLLPSPPTAPDRT 60
OY 61 AVATASRLGPVYLLEPBEGRAYRALHCPGTETCYCKYVPQBALAVLEPYRLPPHKIV 120
DB 61 AVATASRLGPVYLLEPBEGRAYRALHCPGTETCYCKYVPQBALAVLEPYRLPPHKIV 120
OY 121 ARPTVLAGTQLLYAFPTRTHGDMSLVSRHRIPEPAAVLFRQMATALAHCHQGLVL 180
DB 121 ARPTVLAGTQLLYAFPTRTHGDMSLVSRHRIPEPAAVLFRQMATALAHCHQGLVL 180
OY 181 RDLKLCRFVPAFRERKKVLLENLSDCVLTGPDSDLMQACPAVVGPEILSSRASYSK 240
DB 181 RDLKLCRFVPAFRERKKVLLENLSDCVLTGPDSDLMQACPAVVGPEILSSRASYSK 240
OY 241 AADVWSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPARCLVRCLLR 300
DB 241 AADVWSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPARCLVRCLLR 300
OY 301 ABRLTATGILLHPMLRQDPMPLAFTRSHLMEAAQVVDGLGDEAREBEGDREVVLYG 358
DB 301 ABRLTATGILLHPMLRQDPMPLAFTRSHLMEAAQVVDGLGDEAREBEGDREVVLYG 358
RESULT 2
AAU038908 ID AAU038908 standard; Protein; 358 AA.
XX
XX AAU038908;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 2053.
XX
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemostatic;
XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX

OS Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (H58E-) H58E INC.
XX
XX Tang YT, Liu C, Abundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao Q, Zhou P, Goodrich R, Demanac RT;
XX
XX MPI; 2001-442253/47.
XX N-PSDB; AA158064.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 3; SEQ ID NO 2053; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAU038642-AAU042213) with noctropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 358 AA;
SQ
Query Match 99.6%; Score 1884; DB 22; Length 358;
Best Local Similarity 99.4%; Pred. No. 3-9e-171;
Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 MRATPLAAPAGSLSKKRLLELDNDLTERPVOKRARSQPRLPCLLPSPPTAPDRT 60
DB 1 MRATPLAAPAGSLSKKRLLELDNDLTERPVOKRARSQPRLPCLLPSPPTAPDRT 60
OY 61 AVATASRLGPVYLLEPBEGRAYRALHCPGTETCYCKYVPQBALAVLEPYRLPPHKIV 120
DB 61 AVATASRLGPVYLLEPBEGRAYRALHCPGTETCYCKYVPQBALAVLEPYRLPPHKIV 120
OY 121 ARPTVLAGTQLLYAFPTRTHGDMSLVSRHRIPEPAAVLFRQMATALAHCHQGLVL 180
DB 121 ARPTVLAGTQLLYAFPTRTHGDMSLVSRHRIPEPAAVLFRQMATALAHCHQGLVL 180
OY 181 RDLKLCRFVPAFRERKKVLLENLSDCVLTGPDSDLMQACPAVVGPEILSSRASYSK 240
DB 181 RDLKLCRFVPAFRERKKVLLENLSDCVLTGPDSDLMQACPAVVGPEILSSRASYSK 240
OY 241 AADVWSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPARCLVRCLLR 300
DB 241 AADVWSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPARCLVRCLLR 300

QY 301 AERLTATGILHPMLRQDPMPLAPTRSHLWMAAQQVPPDGLGDEARBBSGDRREVLYG 358
 DB 301 AERLTATGILHPMLRQDPMPLAPTRSHLWMAAQQVPPDGLGDEARBBSGDRREVLYG 358

RESULT 3

ID AAB20326 standard; Protein; 358 AA.

AC AAB20326;

DT 29-MAY-2001 (first entry)

DE Human protein phosphatase and kinase protein-5.

KW Protein phosphatase and kinase protein; PPHK-5; human;

KW gastrointestinal disorder; immune system disorder;

KW neurological disorder; cell proliferative disorder; cancer;

OS diagnosis; therapy.

XX Homo sapiens.

XX Key

FT Region

FT Domain

FT Domain

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

CC The present sequence is that of novel human protein phosphatase
 CC and kinase protein PPHK-5, as predicted from incyte clone ID No.
 CC 1271505CB1 (see AAF30480). Tissues that express PPHK-5 (as a
 CC fraction of total tissues expressing PPHK-5) include reproductive
 CC (0.288), gastrointestinal (0.212) and hematopoietic or immune
 CC (0.192). Diseases or conditions associated with tissues expressing
 CC PPHK-5 (as a fraction of total tissues expressing PPHK-5) include
 CC cancer (0.577), inflammation or trauma (0.327) and cell
 CC proliferation (0.308). The encoded protein shows homology to rat
 CC kinase. The invention provides human PPHK-1 to -11 polypeptides
 CC (see AAB20322-32) and polynucleotides (see AAF30476-86). It also
 CC provides expression vectors, host cells, antibodies, agonists and
 CC antagonists, as well as methods for diagnosing, treating or
 CC preventing disorders associated with expression of PPHK, including
 CC gastrointestinal disorders, immune system disorders, neurological
 CC disorders and cell proliferative disorders, including cancer.

SQ Sequence 358 AA;

Query Match 99.6%; Score 1884; DB 22; Length 358;

Best Local Similarity 99.4%; Pred. No. 3.9e-171; Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKRLSLDDNLDTERPVQKRSRSGPRLPCLPLSPPTAPDRAT 60
 DB 1 MRATPLAAPAGSLSRKRLSLDDNLDTERPVQKRSRSGPRLPCLPLSPPTAPDRAT 60
 QY 61 AVATASRLGPPVLLPEBEGRAAYRALACPTGTBYTCKYYPVQBALAVLBPYARLPKHIV 120
 DB 61 AVATASRLGPPVLLPEBEGRAAYRALACPTGTBYTCKYYPVQBALAVLBPYARLPKHIV 120
 QY 121 ARPTVLAGTQLLVAFPTRTGDMHSLVRSRRIRPEPAVLPFQMATRLAHCHQGLVL 180
 DB 121 ARPTVLAGTQLLVAFPTRTGDMHSLVRSRRIRPEPAVLPFQMATRLAHCHQGLVL 180
 QY 181 RDLCRCFPVADREKRLVLSNLEDSCVLTGPDSDLMKHCAPAVGPRISSRSYSCK 240
 DB 181 RDLCRCFPVADREKRLVLSNLEDSCVLTGPDSDLMKHCAPAVGPRISSRSYSCK 240
 QY 241 AADVSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVCLRRBP 300
 DB 241 AADVSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVCLRRBP 300
 QY 301 AERLTATGILHPMLRQDPMPLAPTRSHLWMAAQQVPPDGLGDEARBBSGDRREVLYG 358
 DB 301 AERLTATGILHPMLRQDPMPLAPTRSHLWMAAQQVPPDGLGDEARBBSGDRREVLYG 358

RESULT 4

ID AAM40694 standard; Protein; 393 AA.

AC AAM40694;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 5625.

KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Abundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J,
XX Zhao Q, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX
XX N-PSDB; AAI59850.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 5625; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM3642-AAM42213) with neurotropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilization of the activities such as: immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening and
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 393 AA;
SQ
Query Match 99.6%; Score 1884; DB 22; Length 393;
Best Local Similarity 99.4%; Pred. NO. 4.4e-171;
Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRATPLAAGSLSRKKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTAPDRAT 60
DB 36 MRATPLAAGSLSRKKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTAPDRAT 95
QY 61 AVATASRLGPPVVLSPBEGGAYRALHCTGTGTCCKYPPVDBALAVLSPYARLPKHIV 120
DB 96 AVATASRLGPPVVLSPBEGGAYRALHCTGTGTCCKYPPVDBALAVLSPYARLPKHIV 155
QY 121 ARPTVLGAGTOLVLPFTRTHGDMSLVRSRRIPEPEAAVLFROMATLALACHOGLVL 180
DB 156 ARPTVLGAGTOLVLPFTRTHGDMSLVRSRRIPEPEAAVLFROMATLALACHOGLVL 215
QY 181 RDLKLCRFVADRRKKVLLENLBDSCVLTGPDDSLMDGACPAVYGPPEILSSRSASGK 240
DB 216 RDLKLCRFVADRRKKVLLENLBDSCVLTGPDDSLMDGACPAVYGPPEILSSRSASGK 275
QY 241 AADVWSLGVALLFTMLAGHYPPFODSEPVLLFGKIRGAYALPAGLSAPARCIVRLRRRP 300
DB 276 AADVWSLGVALLFTMLAGHYPPFODSEPVLLFGKIRGAYALPAGLSAPARCIVRLRRRP 335
QY 301 AERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVDGLDEARBEEDREVVLYG 358
DB 336 AERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVDGLDEARBEEDREVVLYG 393

ID ABB80976 standard; Protein; 360 AA.
XX
XX ABB80976;
AC
XX 21-OCT-2002 (first entry)
DT
XX
XX Human tribbles homologue-3 (htrb-3) polypeptide encoding DNA.
DB
XX
XX Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
XX antirheumatic; antiarthritic; antidiabetic; antiproliferative; osteoporotic;
XX ophthalmological; cardiant; cytoskeletal; haemostatic; immunosuppressive;
XX tumour necrosis factor; TNF; htrb-3.
XX
XX Homo sapiens.
XX
XX WO200253743-A2.
XX
XX 11-JUL-2002.
XX
XX 08-JAN-2002; 2002WO-US00070.
XX
XX 08-JAN-2001; 2001US-260294P.
XX
XX (INTE-) INTERLEUKIN GENETICS INC.
XX
XX Dower S, Quantstrom E, Kles-Toth E;
XX
XX WPI; 2002-590635/63.
XX
XX N-PSDB; ABB86479.
XX
XX Novel isolated human tribbles homologue-1 polypeptide for inhibiting
XX AP-1-mediated inflammatory signal in a cell, and activating
XX ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
XX
XX Example 10; Fig 11B; 131pp; English.
XX
XX The invention provides an isolated human tribbles homologue-1 (htrb-1,
XX also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
XX htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
XX signal in a cell. The polypeptide employed in the method is preferably
XX htrb-1, htrb-1 N C, htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N C, or
XX htrb-3 N C. It is also useful for providing htrb agonist activity for
XX activating an ERK-mediated signal e.g. AP-1-mediated gene activation
XX signal, an estrogen receptor-mediated gene activation signal, an
XX fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
XX in a cell. Htrb modulators are useful for modulating AP-1 mediated
XX inflammatory signal in a cell such as tumor necrosis factor (TNF)
XX induced inflammatory signal, or an interleukin induced inflammatory
XX signal. htrb proteins are useful in screening assays, predictive medicine
XX and in therapeutics or prophylactics. The htrb proteins are useful for
XX screening compounds e.g. for treating and/or preventing diseases caused
XX by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
XX psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
XX and cancers. The htrb therapeutics are useful for antagonizing
XX interleukin-1 dependent disorders of human placenta, intraventricular
XX hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
XX and inflammation or autoimmune disorders. The present sequence represents
XX the htrb-3 polypeptide.
XX
XX Sequence 360 AA;
SQ
Query Match 98.9%; Score 1870; DB 23; Length 360;
Best Local Similarity 98.6%; Pred. NO. 8.5e-170;
Matches 355; Conservative 3; Mismatches 0; Indels 2; Gaps 1;
QY 1 MRATPLAAGSLSRKKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTAPDRAT 60
DB 1 MRATPLAAGSLSRKKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTAPDRAT 60
QY 61 AVATASRLGPPVVLSPBEGGAYRALHCTGTGTCCKYPPVDBALAVLSPYARLPKHIV 120

Db 61 AVATASRLGPPVLLPEBEGRAYQALHCPGTETCTCKYYPVOBALAVLEPYARLPKHV 120

Qy 121 ARPTVLAGTQLVAFTRTHGDMHSLVSRSHRIPEPEAAVLFRQMATLAHCHQGLVL 180

Db 121 ARPTVLAGTQLVAFTRTHGDMHSLVSRSHRIPEPEAAVLFRQMATLAHCHQGLVL 180

Qy 181 RDLKLCRFVFAADRRKKLVLENLEDSCVLTGPDLSLMDKHAACPAYVGPBILSSRASYS 240

Db 181 RDLKLCRFVFAADRRKKLVLENLEDSCVLTGPDLSLMDKHAACPAYVGPBILSSRASYS 240

Qy 239 GKADVMSLGVALFTMLAGHYFPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRR 298

Db 241 GKADVMSLGVALFTMLAGHYFPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRR 300

Qy 299 EPARLTATGTGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358

Db 301 EPARLTATGTGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 360

RESULT 6
AA69157
ID AAY69157 standard; Peptide; 360 AA.

XX AC AAY69157;
XX DT 30-MAY-2000 (first entry)
XX DE Peptidea JJ503-XS comprising domains VIA to XI of a protein kinase.
XX KM Kinase activity; molecular weight marker; isoelectric focusing marker;
XX OS Peptide fragmentation control; cellular signal transduction.
XX OS Homo sapiens.
XX PN WO200008180-A2.
XX PD 17-FEB-2000.

XX PF 03-AUG-1999; 99WO-US17630.
XX PR 04-AUG-1998; 98US-0095270.
XX PR 11-SEP-1998; 98US-0099972.
XX PA (IMMV) IMMUNEX CORP.
XX PI Virca GD, Bird TA, Anderson DM, Marken JS;
XX DR WPI; 2000-195584/17.
XX DR N-PSDB; AA261155.

XX PT New human kinase polypeptides and polynucleotides used as molecular
XX weight markers and as controls for peptide fragmentation -
XX PS Claim 2; Page 10; 60pp; English.

XX CC The present sequence represents a partial polypeptide which has kinase
XX activity. The kinase polynucleotides can be used to express the
XX polypeptides, and as probes to identify nucleic acids encoding
XX proteins having kinase activity. The kinase polypeptides and
XX fragmented polypeptides are used as molecular weight and isoelectric
XX focusing markers, and as controls for peptide fragmentation. They also
XX have a number of therapeutic uses as kinases play a central role in
XX cellular signal transduction. The polypeptides could also be used to
XX identify binding partner proteins. The polypeptides can also be used as
XX a reagent to identify any proteins that the polypeptides regulate, and
XX CC proteins with which it might interact. The polypeptides may also be
XX used for preparation of antibodies. The antibodies can be used in
XX assays to detect the presence of the protein, and to purify the protein
XX by immunofluorescence chromatography.

XX Sequence 360 AA;

Query Match 88.6%; Score 1675; DB 21; Length 360;

Best Local Similarity 99.7%; Pred. No. 3.5e-151;
Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKARSQPQPLPCLPLSPPTADPAT 60

Db 43 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKARSQPQPLPCLPLSPPTADPAT 102

Qy 61 AVATASRLGPPVLLPEBEGRAYQALHCPGTETCTCKYYPVOBALAVLEPYARLPKHV 120

Db 103 AVATASRLGPPVLLPEBEGRAYQALHCPGTETCTCKYYPVOBALAVLEPYARLPKHV 162

Qy 121 ARPTVLAGTQLVAFTRTHGDMHSLVSRSHRIPEPEAAVLFRQMATLAHCHQGLVL 180

Db 163 ARPTVLAGTQLVAFTRTHGDMHSLVSRSHRIPEPEAAVLFRQMATLAHCHQGLVL 222

Qy 181 RDLKLCRFVFAADRRKKLVLENLEDSCVLTGPDLSLMDKHAACPAYVGPBILSSRASYS 240

Db 223 RDLKLCRFVFAADRRKKLVLENLEDSCVLTGPDLSLMDKHAACPAYVGPBILSSRASYS 282

Qy 241 AADVMSLGVALFTMLAGHYFPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRR 300

Db 283 AADVMSLGVALFTMLAGHYFPQDSBPVLLFGKIRGAYALPAGLSAPARCLVRCILRR 342

Qy 301 AERLTATGTGILLHPMLRQD 318

Db 343 AERLTATGTGILLHPMLRQD 360

RESULT 7
AAB85791
ID AAB85791 standard; Protein; 323 AA.

XX AC AAB85791;
XX DT 29-OCT-2001 (first entry)
XX DE Human kinase PKIN-10.

XX KM PKIN; kinase; cytosolic; immunosuppressive; immunostimulant; human;
XX antileukocytotoxic; cardiant; gene therapy; antisense therapy.

XX OS Homo sapiens.
XX PN WO200160991-A2.
XX PD 23-AUG-2001.

XX PF 16-FEB-2001; 2001WO-US05240.
XX PR 17-FEB-2000; 2000US-0183682.
XX PR 02-MAR-2000; 2000US-0186559.
XX PR 09-MAR-2000; 2000US-0186506.
XX PR 17-MAR-2000; 2000US-0189998.
XX PR 30-MAR-2000; 2000US-0193851.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Tang YT, Buford N, Gandhi AR, Patterson C, Khan FA, Yue H;
XX PI Hafealia A, Shih LY, Tribouley CM, Yao MG, Burrill JD, Marcus GA;
XX PI Zingler KA, Lu DM, Bandman O, Policky JL, Griffin JA, Thornton M,
XX PI Nguyen DB, Lal P, Walsh RT;
XX DR WPI; 2001-514771/56.
XX DR N-PSDB; AAH76218.

XX PT isolated human kinase polypeptides useful in the diagnosis, treatment
XX and prevention of cancer, immune disorders and disorders affecting
XX growth and development -
XX PS Claim 1; Page 115; 126pp; English.

XX CC The invention provides human kinases (PKIN) and polynucleotides encoding
XX PKIN. The PKIN polypeptides can be expressed using standard recombinant

CC methodology. The PKIN polypeptides, polynucleotides, modulators and
CC specific antibodies are useful in the diagnosis, treatment and prevention
CC of cancer, immune disorders, disorders affecting growth and development,
CC atherosclerosis, and other cardiovascular diseases, and lipid disorders
CC and in the assessment of the effects of exogenous compounds on the
CC expression of nucleic acid sequences of human kinases. The present
CC sequence represents a human PKIN-10 polypeptide.

5Q Sequence 323 AA;

Query Match	83.7%	Score 1583.5	DB 22	Length 323
Best Local Similarity	86.0%	Pred. No. 2e-142		
Matches 308, Conservative	4	Mismatches 11	Indels 35	Gaps 2

QY	1	NRATGLAIPAGLSKKRRL	ELDNDL	DTES	PVOKRA	SGOPRL	PCLL	PLSP	TAD	DRAT	60																									
Db	1	NRATTLAASAGLSKKRRL	ELDNDL	DTET	PVOKRA	SGOPRL	PCLL	PLSP	TAD	DRAT	60																									
QY	61	AVATASRLGPVYLLEP	BEGG	RAYALH	CFGT	BYTCTKY	PVQDAL	AVLEP	YARL	PEHKIV	120																									
Db	61	AVATASRLGPVYLLEP	BEGG	RAYALH	CFGT	EYTC	KVPVQDAL	AVLEP	YARL	PEDKIV	120																									
QY	121	ARPTAVLAGTOLLV	PFRT	THDD	MSLV	SRRI	PE	BA	AV	PRQW	ATLACHQ	GLVL	180																							
Db	121	AMPTGGLAGTOLLV	PFRT	THDD	MR	LIGH	-----	T	C	A	N	C	-----	156																						
QY	181	RDULKCRFV	ADRE	NKKVL	ENL	EDS	CVL	TG	PD	DS	LM	D	R	A	C	P	AYV	GP	IL	S	R	A	S	Y	S	G	K	240								
Db	157	-----	DOT	R	K	V	L	EN	L	E	B	S	CVL	TG	PD	DS	LM	D	R	A	C	P	AYV	GP	IL	S	R	A	S	Y	S	G	K	205		
QY	241	AADVMSLGVAL	PTML	AGHY	P	FO	D	S	EVLL	F	G	K	I	R	G	A	Y	A	L	P	G	L	S	A	P	A	C	L	R	C	L	R	R	E	P	300
Db	206	AAAVMSLGVAL	PTML	AGHY	P	FO	D	S	EVLL	F	G	K	I	R	G	A	Y	A	L	P	G	L	S	A	P	A	C	L	R	C	L	R	R	E	P	265
QY	301	ABRLTATGILL	HPWL	R	OD	P	M	L	A	P	T	R	SH	ME	A	Q	V	P	D	G	L	D	E	A	B	E	G	D	E	V	L	Y	G	358		
Db	266	ABRLTATGILL	HPWL	R	OD	P	M	L	A	P	T	R	SH	ME	A	Q	V	P	D	G	L	D	E	A	B	E	G	D	E	V	L	Y	G	323		

RESULT 8
AAB43657
ID AAB43657 standard; protein; 233 AA.

AC AAB43657

DT 08-FEB-2001 (first entry)

DB	Human cancer associated protein sequence	SEQ ID NO:1102.
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

KM Human, cancer associated gene; cancer antigen; detection; cancer;
 KM diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KM antidiabetic; antistatic; antirheumatic; antiarthritic; antiviral;
 KM antinflammatory; antihypoid; antiallergic; antidiabetic; cardiant;
 KM dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KM vasotropic; antipneumatic; angiogenic; gene therapy; inflammation;
 KM immune disorder; hematopoietic cell disorder; autoimmune disorder;
 KM allergic reaction; graft versus host disease; organ rejection;
 KM hemostatic; thrombolytic; cardiovascular; infection;
 KM neurological disease; drug screening.

Homo sapiens.

PN W0200055350-A1.

PD 21-8EP-2000.

08-MAR-2000; 2000WO-US05882.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rozen CA, Ruben SM;

XX WPI, 2000-587533/55.
DR N-PSDB, AAC77866.
DR
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
PT
XX
PS Claim 11, Page 1707-1708, 2352pp; English.

CC AAC776072 to AAC78444 encode the human cancer associated proteins given
CC in AAB33398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytotoxic; proliferative; vlnary; immunomodulator;
CC antidiabetic; antiaesthetic; antihemetic; antarthritic;
CC antinflammatory; antitumor; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
CC neurotropic; vasotropic; antiproliferative and angiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC The polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.

SQ Sequence 233 AA;

Query Match	58.3%	Score 1102;	DB 21;	Length 233;
Best Local Similarity	100.0%;	Pred. No. 9;	2e-97;	
Matches 208;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	151	RHRIPERPAALVLFQMAATALAHCHQHGVLVDLKLCPVPADRSKKLYLKNLEDSCVLT	210
Db	26	RHRIPERPAALVLFQMAATALAHCHQHGVLVDLKLCPVPADRSKKLYLKNLEDSCVLT	85
QY	211	GPDDSLMDKACAPAYGPEILSSRASYSGKADVMSLGVALLFTMLAGHPFQDSEPVLLP	270
Db	86	GPDDSLMDKACAPAYGPEILSSRASYSGKADVMSLGVALLFTMLAGHPFQDSEPVLLP	145
QY	271	GKIRGAYALPAGLSAPARCLVRCILRRBPABRLTATGILLHPWLRQDPMPLAPTRSHLM	330
Db	146	GKIRGAYALPAGLSAPARCLVRCILRRBPABRLTATGILLHPWLRQDPMPLAPTRSHLM	205
QY	331	EAQGVPPDGLGLDSEARREEGDRBVLTYG	358
Db	206	EAQGVPPDGLGLDSEARREEGDRBVLTYG	233

ID ABB06093 standard; Protein; 278 AA.

AC ABB06093

DT 10-MAY-2002 (first entry)

DB Human NS protein sequence SEQ ID NO:185.

Human, osteopathic, synaesthetical, neuroprotective;
KM antihemmatic; antiarthritic; antipsoriatic; ophthalmological, anti-HIV;
KM vasotropic, antisterioelectroic; antinflammatory, dermatological;
KM anorectic, muscular, antilethargic, cardiovascular, anticoagulant;
KM antibrinolytic, hypotension, antiaesthetic; immunomodulator; cardiac;
KM anticonvulsant; antidiabetic; tranquiliser; antidepressant; neuroleptic;
KM gastrointestinal, vitruvian; antiqualis; cerebroprotective, nootropic;
KM contraceptive; vaccine; gene therapy; cancer; osteoporosis; dysomnia;
KM endometriosis; degenerative disease; multiple sclerosis; psoriasis;

KM rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KM inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KM infertility; cardiovascular disease; coagulation disease; hypertension;
 KM ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KM diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KM gastric ulcer; Alzheimer's disease.

XX Homo sapiens.

OS WO200206315-A2.

XX 24-JAN-2002.

XX 17-JUL-2001; 2001WO-1100653.

XX 18-JUL-2000; 2000IL-0137345.

XX 15-DEC-2000; 2000IL-0140354.

XX (COMP-) COMPUGEN LTD.

XX Mintz L, Freilich S, Bernstein J,

XX WPI; 2002-155037/20.

XX N-PSDB; ABL39747.

XX One hundred and twenty eight novel nucleic acid sequences, useful for

XX treating and diagnosing e.g. cancer, asthma and Alzheimer's -

XX Claim 6; Page 213-214; 290pp; English.

XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antineumatic, antiarthritic, antipsoriatic, ophthalmological, vitruide,
 CC vasorectic, antiarteriosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, anti-HIV, antifertility, cardiovascular,
 CC anticonvulsant, antifibrinolytic, hypotension, antiaesthetic, cardiac,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
 CC antidepressant, gastroenteric, antileptic, cerebroprotective,
 CC neurotropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive.

XX Sequence 278 AA;

XX Query Match 58.0%; Score 1097; DB 23; Length 278;

XX Best Local Similarity 99.1%; Pred. No. 3.5e-96;

XX Matches 209; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

XX 1 MRATPLAAPAGSGSRKKRLDNLDTERRVQKARGSPRPLPCLPSPAPRRAT 60

XX 29 MRATPLAAPAGSGSRKKRLDNLDTERRVQKARGSPRPLPCLPSPAPRRAT 88

XX 61 AVATTAARLGPVYLPEPEGRAVALHCPGTGTCVYVORALVLEPYARLPKHV 120

XX 89 AVATTAARLGPVYLPEPEGRAVALHCPGTGTCVYVORALVLEPYARLPKHV 148

XX 121 ARPTVLAAGTOLLVAFRTTHGDMHSLVRGRHRIPEBEAAVLFROMATALAHCHQGLV 180

XX 149 ARPTVLAAGTOLLVAFRTTHGDMHSLVRGRHRIPEBEAAVLFROMATALAHCHQGLV 208

XX 181 RDLKLCRPVADDERKKLVLENEDGCVLTG 211

XX 209 RDLKLCRPVADDERKKLVLENEDGCVLTG 239

RESULT 10

ABB06108

ID ABB06108 standard; Protein; 278 AA.

XX ABB06108;

XX 10-MAY-2002 (first entry)

XX Human NS protein sequence SEQ ID NO:200.

XX Human; cytostatic; osteopathic; gynaecological; neuroprotective;

XX antineumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;

XX vasorectic; antiarteriosclerotic; antiinflammatory; dermatological;

XX anorectic; muscular; antifertility; cardiovascular; anticonvulsant;

XX anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;

XX gastroenteric; vitruide; antiulcer; cerebroprotective; neurotropic;

XX endometriosis; degenerative disease; multiple sclerosis; psoriasis;

XX rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;

XX inflammation; skin disorder; obesity; muscular dystrophy; AIDS;

XX infertility; cardiovascular disease; coagulation disease; hypertension;

XX ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;

XX diabetes; anxiety; depression; schizophrenia; viral disease; stroke;

XX gastric ulcer; Alzheimer's disease.

XX (COMP-) COMPUGEN LTD.

XX Mintz L, Freilich S, Bernstein J,

XX WPI; 2002-155037/20.

XX N-PSDB; ABL39747.

XX One hundred and twenty eight novel nucleic acid sequences, useful for

XX treating and diagnosing e.g. cancer, asthma and Alzheimer's -

XX Claim 6; Page 231-232; 290pp; English.

XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antineumatic, antiarthritic, antipsoriatic, ophthalmological, vitruide,
 CC vasorectic, antiarteriosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, anti-HIV, antifertility, cardiovascular,
 CC anticonvulsant, antifibrinolytic, hypotension, antiaesthetic, cardiac,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
 CC antidepressant, gastroenteric, antileptic, cerebroprotective,
 CC neurotropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive.

XX Sequence 278 AA;

Query Match 59.04; Score 1097; DB 23; Length 278;
 Best Local Similarity 99.14; Pred. No. 3.5e-96;
 Matches 209; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKKRLDLDLDTBRPVQKARSGGPPRLPCLPLSPAPRAT 60
 |||||
 DB 29 MRATPLAAPAGSLSRKKRLDLDLDTBRPVQKARSGGPPRLPCLPLSPAPRAT 88
 |||||
 QY 61 AVATASRLGPVYLPEPEGGRAYRALHCPGTGTYCKVYQVQALAVLEPYARLPKHV 120
 |||||
 DB 89 AVATASRLGPVYLPEPEGGRAYRALHCPGTGTYCKVYQVQALAVLEPYARLPKHV 148
 |||||
 QY 121 ARPEVLATGTLVAFPTRTGDMSLVRSNRIRPEBAAVLFROMATALACHQGLVL 180
 |||||
 DB 149 ARPEVLATGTLVAFPTRTGDMSLVRSNRIRPEBAAVLFROMATALACHQGLVL 208
 |||||
 QY 181 RDLKLCRFVADRERKKLVLENLEDSCLVTG 211
 |||||
 DB 209 RDLKLCRFVADRERKKLVLENLEDSCLVTG 239
 |||||

RESULT 11
 ABB80975
 ID ABB80975 standard; Protein; 372 AA.
 AC ABB80975;
 DT 21-OCT-2002 (first entry)
 XX
 XX Human tribbles homologue-1 (htrb-1) polypeptide.
 DE
 XX Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; Ap-1;
 KW antitumoural; antiarthritic; antidiabetic; antiparoxysmic; osteopathic;
 KW ophthalmological; cardiac; cytosolic; haemostatic; immunosuppressive;
 KW antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
 KW tumour necrosis factor; TNF.
 KM
 XX Homo sapiens.
 OS
 XX MO200253743-A2.
 PN
 XX 11-JUL-2002.
 PD
 XX 08-JAN-2002; 2002MO-US00070.
 PP
 XX 08-JAN-2001; 2001US-260294P.
 PR
 XX (INTB-) INTERLUKIN GENETICS INC.
 PA
 XX Dower S, Quanstrom E, Kiss-Toth E;
 PI WPI; 2002-590635/63.
 DR N-PSDB; ABB86478.
 DT
 XX Novel isolated human tribbles homologue-1 polypeptide for inhibiting
 PT AP-1-mediated inflammatory signal in a cell, and activating
 PT ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
 PT
 XX Claim 21; Fig 10B; 131pp; English.
 PS
 XX The invention provides an isolated human tribbles homologue-1 (htrb-1,
 CC also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
 CC htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
 CC signal in a cell. The polypeptide employed in the method is preferably
 CC htrb-1, htrb-1 N htrb-1 C, htrb-1 C, htrb-3, htrb-3 N htrb-3 C, or
 CC htrb-3 N C. It is also useful for providing htrb agonist activity for
 CC activating an ERK-mediated signal e.g. AP-1-mediated gene activation
 CC signal, an estrogen receptor-mediated gene activation signal, an
 CC fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
 CC in a cell. Htrb modulators are useful for modulating AP-1 mediated
 CC inflammatory signal in a cell such as tumor necrosis factor (TNF)
 CC induced inflammatory signal, or an interleukin induced inflammatory

CC signal. htrb proteins are useful in screening assays, predictive medicine
 CC and in therapeutics or prophylactics. The htrb proteins are useful for
 CC screening compounds e.g. for treating and/or preventing diseases caused
 CC by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
 CC psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
 CC and cancer. The htrb therapeutics are useful for antagonizing
 CC interleukin-1 dependent disorders of human placenta, intraventricular
 CC hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
 CC and inflammation or autoimmune disorders. The present sequence represents
 CC the htrb-1 polypeptide.

Sequence 372 AA;

Query Match 39.44; Score 744.5; DB 23; Length 372;
 Best Local Similarity 46.84; Pred. No. 2.3e-62;
 Matches 162; Conservative 47; Mismatches 126; Indels 11; Gaps 6;

QY 2 RATPLAAPA-GSLSRKKRLDLD--NIDTRPVQKARSGP--QPLPCLPLSPPTA 55
 |||||
 DB 17 RGPALFPATRGVPAKRLDADDAVAACPCPLSBCSPDYLSPGSPC-SFQPPAA 75
 |||||
 QY 56 PDRTAVATA---SLGPVYLPEPEGGRAYRALHCPGTGTYCKVYQVQALAVLEPYA 112
 |||||
 DB 76 PGAGGGSGSAPGSRINDVLLPLAERHVSRLCTITGRBLCKVPIKHYDKIRPYI 135
 |||||
 QY 113 RLPPHKIVARPTVLAGTOLLVAFPTRTGDMSLVRSNRIRPEBAAVLFROMATALAH 172
 |||||
 DB 136 QLPSSHNTIGVIVELGETRAYVFEKSPFQDMISYVSRKRLREBAARLPKQVSAVAH 195
 |||||
 QY 173 CHQGLVLRDLKLCRFVADRERKKLVLENLEDSCLVTGDDSLMDGACPAYVGPRLS 232
 |||||
 DB 196 CHQSAIVLGLKLRKVFSTBERTQLRLSLBQTHIMKGBDDLSDRGCPAYVSPBILN 255
 |||||
 QY 233 SRASVSGKADVMSLGVALLTTLAGHYFPQDSPPVLLFGIRGAYALPGLSAPACLV 292
 |||||
 DB 256 TTGTYSKADVMSLGVMLTTLVGRVFPDSDPSALFSIKRGQFCIPRHISKACLI 315
 |||||
 QY 293 RCLLRREPARLTATGILLHPLRQDPMPLAPTRSHLMAQVVPD 338
 |||||
 DB 316 RSLLRREPSRLTAPRILHPLMFESVLEP-GYIDSLRGTSQDQIVPE 360
 |||||

RESULT 12
 AAU28135
 ID AAU28135 standard; Protein; 269 AA.
 AC AAU28135;
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secretory protein, Seq ID No 304.
 XX
 XX Human; secreted protein; arthritic; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 KM
 XX Homo sapiens.
 OS
 XX MO20016689-A2.
 PN
 XX 13-SEP-2001.
 PD
 XX 05-MAR-2001; 2001MO-US04942.
 PP
 XX 07-MAR-2000; 2000US-0517905.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.

PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0655363.
 PR 20-OCT-2000; 2000US-0693267.
 PA (HYSB-) HYSBQ INC.
 PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 DR WPI, 2001-589934/66.
 DR N-PSDB; AAS45035.
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 XX
 XX
 XX Example 4; SEQ ID No 304; 107bp; English.
 CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of hematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorythms or circadian cycles of rhythm,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.
 CC
 XX
 XX Sequence 269 AA;
 SQ
 Query Match 36.0%; Score 681.5; DB 22; Length 269;
 Best Local Similarity 51.8%; Pred. No. 1.5e-56;
 Matches 133; Conservative 39; Mismatches 79; Indels 7; Gaps 2;
 QY 88 CP-----TGTETCTCKYPPQDALVALPEPAKLPKHGKAPPEVLAGTOLLAFPTRT 141
 DB 2 CPGRCASTLGRARRVKCFPIKHQYODKIRPYIQLPSHNSINGIYEIVIGRKAVYFFPKDR 61
 QY 142 GDMHSLVSRHRIPEPEAVLFRQMATLALHCHQHGVLNLDLKLRCFVPADRREKLVLE 201
 DB 62 GDMHSLVSRHRIPEPEAVLFRQMATLALHCHQHGVLNLDLKLRCFVPADRREKLVLE 121
 QY 202 NLEDSCVLTPGDDSLMDKHAACPAYVGBEILSSASYSYGKADWSLGVALLFTMLAGHYPP 261
 DB 122 SLEDTHIMKGEDDALSDKHCAPAYVSPBEILNTGTYSYGKADWSLGVALLFTMLAGHYPP 181
 QY 262 QDSEPTVLLPGKIRGAYALPAGISAPARCVCVRLCLRRREPRLTATGILLHPMLRQDPMP 321
 DB 182 HSDSPALPSPKIRRGQFCIPBHSIPKARCLIRSLRRBSPERLTAPBILHPWFSVLEP 241

QY 322 LAPTRSHMEAAQVVPD 338
 DB 242 -GYIDSEIGTSDQIVPS 257
 RESULT 13
 AAU28323
 ID AAU28323 standard; Protein; 290 AA.
 XX
 AC AAU28323;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 XX Novel human secretory protein, Seq ID No 680.
 XX
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 XX
 OS Homo sapiens.
 XX
 PN W020016689-A2.
 PD 13-SEP-2001.
 XX
 PD 05-MAR-2001; 2001MO-US04942.
 XX
 PR 07-MAR-2000; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0655363.
 PR 20-OCT-2000; 2000US-0693267.
 XX
 XX (HYSB-) HYSBQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX
 XX WPI, 2001-589934/66.
 DR N-PSDB; AAS45223.
 DR
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 PT
 XX
 XX Example 2; SEQ ID No 680; 107bp; English.
 CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of hematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,

CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (1) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues. Various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (1) affects biorythms or circadian cycles of rhythms,
 CC fertility, metabolism, carbohydrate, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC anabolic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.

XX Sequence 290 AA:

Query Match 34.9%; Score 659.5; DB 22; Length 290;

Best Local Similarity 51.2%; Pred. No. 2.1e-54;

Matches 127; Conservative 42; Mismatches 76; Indels 1; Gaps 1;

QY 91 GTEVTCCKYRPOEALVLEPYARLPKHVAPFEVLGTLVAFPTRTGDMHSLVRS 150
 DB 22 GPPSLQVFPKHODKIRPYQLPSHSNITGIVLIGETKAYVFEKDFQDMHSYVRS 81
 QY 151 RHRITPEPAVLFROMATALAHCHQGLVRLDKLCRFVADRRKKVLBNLBDSCVLT 210
 DB 82 RKRLREBEAARLFKQIVSAVNHCHQSAVLGDLKLRKVFSTREBTOURLSELDETHLMK 141
 QY 211 GPDSLSMDKACPAVYGPBEILSSRASYSKKAADVMSLGVALFTMLAGHYPPDSRPVLLF 270
 DB 142 GEDALSDKHGCPAVYSPBEILNTGTYSKKAADVMTGVMLYTLVGRYPFHDSPSALF 201
 QY 271 KIRRGAVLALPAGSAPARCIVRCILRRBPARRLTATGILLHPMLRQDPMPLAPRSHLM 330
 DB 202 SKIRRGQCFIEHISPKARCLIRSLRRBPERRLTAPBILHPWESVLEP-GYIDSEIG 260
 QY 331 EAAQVVPD 338
 DB 261 TSDQIVPE 268

RESULT 14

ID ABP96856 standard; Protein; 206 AA.

XX ABP96856;

DT 11-JUN-2003 (first entry)

DB Human C8FW protein SEQ ID NO:2.

KW Human; phosphoprotein regulating mitogenic pathway gene; cytosolic;

KW vaccine; cancer; C8FW.

OS Homo sapiens.

XX MO2003018769-A2.

PD 06-MAR-2003.

XX 27-AUG-2002; 2002WO-US27187.

XX 27-AUG-2001; 2001US-314655P.

XX 31-OCT-2001; 2001US-330797P.

XX (TULDA-) TULARIK INC.

XX L1 J;

XX WPI; 2003-278653/27.

XX N-B8DB; ACC45125; ACC45126.

XX Diagnosing, preventing, treating and prognosticating a cancer in a
 PT mammal by detecting and measuring the C8FW gene copy number and/or C8FW
 PT level, useful for breast, brain, lung, colon, ovarian and/or prostate
 PT cancers -
 XX Claim 16; Page 86; 86pp; English.

XX The present invention describes a method for diagnosing a cancer in a
 CC mammal. The method comprises detecting and measuring the C8FW gene copy
 CC number or level in a biological subject from a region of the mammal that
 CC is suspected to be pre-cancerous or cancerous to generate data for a
 CC test gene copy, and comparing the test gene copy number or level to a
 CC data for a control gene copy number or level, where an amplification of
 CC the gene in the biological subject relative to the control indicates the
 CC presence of a pre-cancerous lesion or cancer in the animal. C8FW has
 CC cytosolic activity, and can be used in vaccines. The method can be used
 CC for the diagnosis, prevention, treatment and prognostication of breast,
 CC colon, lung, brain, prostate and/or ovarian cancer. The present sequence
 CC represents human C8FW from the present invention. C8FW is a
 CC phosphoprotein regulating mitogenic pathway protein.

XX Sequence 206 AA;

Query Match 30.6%; Score 578.5; DB 24; Length 206;

Best Local Similarity 57.9%; Pred. No. 7.1e-47;

Matches 113; Conservative 29; Mismatches 52; Indels 1; Gaps 1;

QY 144 MHSIVSRHRRPEBAVLFROMATALAHCHQGLVRLDKLCRFVADRRKKVLBNL 203
 DB 1 MHSIVSRHRRPEBAVLFROMATALAHCHQGLVRLDKLCRFVADRRKKVLBNL 60
 QY 204 EDSCLTGPDDSLMDKACPAVYGPBEILSSRASYSKKAADVMSLGVALFTMLAGHYPPD 263
 DB 61 EDTHIMKGEDALSDKHGCPAVYSPBEILNTGTYSKKAADVMSLGVMLYTLVGRYPFHD 120
 QY 264 SEPVLLFGKIRGAVLALPAGSAPARCIVRCILRRBPARRLTATGILLHPMLRQDPMPLA 323
 DB 121 SDPSALFSEKIRRGQCFIEHISPKARCLIRSLRRBPERRLTAPBILHPWESVLEP-G 179
 QY 324 PTRSHLEAAQVVPD 338
 DB 180 YIDSEIGTSDQIVPE 194

RESULT 15

ID ABB71379 standard; Protein; 484 AA.

XX ABB71379;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 40929.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PERKE) PE CORP NY.

XX Venter JC, Adams M, Li FWD, Myers BW;

DR WPI; 2001-656860/75.
DR N-PSDB; ABL15482.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 40929; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 484 AA;
Query Match 23.0%; Score 435; DB 22; Length 484;
Best Local Similarity 33.4%; Pred. No. 1.1e-32;
Matches 116; Conservative 45; Mismatches 134; Indels 52; Gaps 8;
QY 31 VQKXASGPOPRLPCLPLSPPTAPDAVATASRLGPLYLLPBEGRAYRALHCP 90
DB 104 IQGRYLISAQPSHISAALAAKTP-ASYRHLVDLTASNL-----RCVDIFT 147
QY 91 GTEYTCAPYVQEL-AVLEPYARLPKHVARPTEVLA-----GTQLL 133
DB 148 GEOFLCRI--VNEPLHVVQRAYFQLOQHDEELRSTIYGHPILRPVHDIPLTKDRITYIL 205
QY 134 YA-----PFRTHGDMHSLVRSRHRIPRPAAVLFRMATALAHCHOGVLRDL 183
DB 206 IAPVQGRDSTGVTGYENHTYIRNAKRLCTEBAIRFQICQYQVCHRNQIILRDL 265
QY 184 KLCRFVADRRKKLVLENLEDSCVLTGPDDSLMDKHAACPAYVGPBILSSRASYSGRAD 243
DB 266 KLRFPYFIDEARTQLQYESLEGSWILDGSDDTLSDKIGCPLYTAPBELLCPQOTYKGRAD 325
QY 244 VMSLGVALFTMLAGHYPRQDSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRREPAER 303
DB 326 MMSLGVALFTMLVGOYPRTEKANCNLITVIRHGNVQIPLTLSKSVRMLLSLRKDYTER 385
QY 304 LTAAGILHPLRQDPMPLAFTRSHLWEAAQVVPDGLDEAREE 350
DB 386 MTASHIFLTPLMRQ----RPFHMYLPVDVEVAED---WSDAREDEG 425

Search completed: January 15, 2004, 14:57:50
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:03:07 ; Search time 41 Seconds

(without alignment)
2253.240 Million cell updates/sec

Title: US-09-909-474D-2

Perfect score: 358
Sequence: 1 MRATPLAAPAGLSRKRLKLE.....GLGLDEAREBEGDREVLYG 358

Scoring table: OLIGO 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	3.9	257	11	Q8R2V8
2	14	3.9	343	4	Q92519
3	14	3.9	343	6	Q28283
4	14	3.9	343	11	Q8K4K3
5	14	3.9	343	11	Q8K017
6	13	3.6	224	4	Q15180
7	13	3.6	364	11	Q9EQJ6
8	13	3.6	372	4	Q96R08
9	13	3.6	372	11	Q9H2Y8
10	13	3.6	372	11	Q8K4K4
11	13	3.6	372	11	Q91W04
12	10	2.8	339	10	Q8H7S4
13	9	2.5	317	16	Q9F3C2
14	9	2.5	562	5	Q18041
15	9	2.5	583	5	Q9X266
16	8	2.2	184	10	Q9LR83

17	8	2.2	190	16	Q92XU5	Q92XU5 rhizobium m
18	8	2.2	249	2	Q9SZ07	Q9SZ07 streptomyce
19	8	2.2	249	16	Q9L184	Q9L184 streptomyce
20	8	2.2	277	4	Q9NS63	Q9NS63 homo sapien
21	8	2.2	325	11	Q9JUD6	Q9JUD6 mus musculu
22	8	2.2	325	11	Q8BNY7	Q8BNY7 mus musculu
23	8	2.2	367	10	Q9AOV3	Q9AOV3 oryza sativ
24	8	2.2	375	16	Q8P930	Q8P930 leptocapira
25	8	2.2	376	4	Q9BYG4	Q9BYG4 homo sapien
26	8	2.2	427	9	Q8SD32	Q8SD32 pseudomonas
27	8	2.2	428	16	Q9RD14	Q9RD14 streptomyce
28	8	2.2	499	11	Q8C0V7	Q8C0V7 mus musculu
29	8	2.2	575	4	Q9H6R5	Q9H6R5 homo sapien
30	8	2.2	575	4	Q8TBX5	Q8TBX5 homo sapien
31	8	2.2	608	12	Q41253	Q41253 rice tagged
32	8	2.2	696	4	Q81VP5	Q81VP5 homo sapien
33	8	2.2	725	16	Q9H2G0	Q9H2G0 pseudomonas
34	8	2.2	813	16	Q8YJW9	Q8YJW9 bruceella me
35	8	2.2	814	16	Q8FY69	Q8FY69 bruceella au
36	8	2.2	934	6	Q9GMD3	Q9GMD3 bos taurus
37	8	2.2	1159	5	Q9VVC9	Q9VVC9 drosophila
38	8	2.2	1715	11	Q9WVS3	Q9WVS3 mus musculu
39	8	2.2	2038	12	Q91H80	Q91H80 cherry necr
40	7	2.0	59	10	Q8GX77	Q8GX77 arabidopsis
41	7	2.0	66	16	Q9Z4Y0	Q9Z4Y0 streptomyce
42	7	2.0	72	16	Q8U8M5	Q8U8M5 agrobacteri
43	7	2.0	76	11	Q8K4U2	Q8K4U2 mus musculu
44	7	2.0	80	16	Q9SP0	Q9SP0 staphylococ
45	7	2.0	88	12	Q8QWS9	Q8QWS9 norwayk vit

ALIGNMENTS

RESULT 1	Q8R2V8	PRELIMINARY:	PRT:	257 AA.
AC	Q8R2V8	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DS	Hypothetical 29.2 kDa protein (Fragment).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxId=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Strasbourg R.I.			
RL	Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL, BC027159; AAH27159.1; .			
DR	InterPro; IPR000719; Prot. kinase.			
DR	Pfam; PF00069; Pkinase; 1.			
DR	ProDom; PD000001; Prot. kinase; 1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
KW	Hypothetical protein; ATP-binding; Transferase.			
FT	NON TER			
FT	SEQUENCE 257 AA; 29232 MW; 84EBACDD4767F51 CRC64;			
Query Match	3.9%; Score 14; DB 11; Length 257;			
Best Local Similarity	100.0%; Pred. No. 5.1e-05;			
Matches	14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	236 SYSGKADVWSIGV 249			
DB	143 SYSGKADVWSIGV 156			
RESULT 2	Q92519	PRELIMINARY:	PRT:	343 AA.
ID	Q92519	01-FEB-1997 (TREMBLrel. 02, Created)		

DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DB G33955 (G33955 protein).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_Taxid=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cancellous bone;
 RA Chou T., Hashimoto J., Takao K., Ochi T., Okubo K., Matsubara K.;
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RA Strauberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, D87119; BAA13250.1; -.
 DR EMBL, BC002637; AAH02637.1; -.
 DR HSP, Q63450; IAO6.
 DR InterPro: IPR000719; Prot. Kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD00001; Prot. Kinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KM ATP-binding; Transferase.
 SQ SEQUENCE 343 AA; 3880 MW; BFB7366DACB84FA CRC64;
 Query Match 3.9%; Score 14; DB 4; Length 343;
 Best Local Similarity 100.0%; Pred. No. 6.7e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 SYSGRADVMSLGV 249
 DB 229 SYSGRADVMSLGV 242
 RESULT 3
 Q28283 PRELIMINARY; PRT; 343 AA.
 ID Q28283;
 AC Q28283;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE C5FW ORF protein.
 GN C5FW ORF.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_Taxid=9615;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE-97067069; PubMed-8910471;
 RA Wilkin P., Savonet V., Radulescu A., Petermans J., Dumont J.E.,
 Maenhaut C.;
 RT Identification and Characterization of Novel Genes Modulated in the
 RT Thyroid of Dogs Treated with Methimazole and Propylthiouracil.;
 RL J. Biol. Chem. 271:28451-28457(1996).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE-98000262; PubMed-9342215;
 RA Wilkin P., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,
 Dumont J.E., Maenhaut C.;
 RT Characterization of a phosphoprotein whose mRNA is regulated by the
 RT mitogenic pathways in dog thyroid cells.;
 RL Eur. J. Biochem. 248:660-669(1997).
 DR EMBL, X99144; CAA67581.1; -.
 DR HSP, Q63450; IAO6.
 DR InterPro: IPR000719; Prot. Kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD00001; Prot. Kinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

KM ATP-binding; Transferase.
 SQ SEQUENCE 343 AA; 38786 MW; BFD1300DACB84FA CRC64;
 Query Match 3.9%; Score 14; DB 6; Length 343;
 Best Local Similarity 100.0%; Pred. No. 6.7e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 SYSGRADVMSLGV 249
 DB 229 SYSGRADVMSLGV 242
 RESULT 4
 Q8K4K3 PRELIMINARY; PRT; 343 AA.
 ID Q8K4K3;
 AC Q8K4K3;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE TRB-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Klags-Toch E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
 Bagstaff S.M., Wylie D.H., Harte M., O'Neill L.A.J., Ovarstrom E.E.,
 Dower S.K.;
 RT "Mammalian homologs of Drosophila tribbles (htb) control mitogen
 RT activated protein kinase signaling."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF358867; AAM45477.1; -.
 DR InterPro: IPR000719; Prot. Kinase.
 DR InterPro: IPR002290; Ser. Thr. Kinase.
 DR InterPro: IPR01245; Tyr. Kinase.
 DR Pfam: PF00069; pkinase; 2.
 DR ProDom: PD00001; Prot. Kinase; 1.
 DR SMART: SM00220; S. TRC; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KM ATP-binding; Transferase.
 SQ SEQUENCE 343 AA; 38758 MW; 0B3965B8B2087D74 CRC64;
 Query Match 3.9%; Score 14; DB 11; Length 343;
 Best Local Similarity 100.0%; Pred. No. 6.7e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 SYSGRADVMSLGV 249
 DB 229 SYSGRADVMSLGV 242
 RESULT 5
 Q8K017 PRELIMINARY; PRT; 343 AA.
 ID Q8K017;
 AC Q8K017;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE TRB-2 (C5FW ORF protein homolog).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strauberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Cerebellum, and Retina;

RX MEDLINE-22354683; PubMed-12466851;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; BC034338; AAH34338.1; -
 DR EMBL; AK044747; BAC32063.1; -
 DR EMBL; AK080064; BAC37820.1; -
 DR EMBL; AK082329; BAC38467.1; -
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR pfam; PF00069; pkinase; 2.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 343 AA; 38772 MW; 9418B7AC19FC23F CRC64;

Query Match 3.6%; Score 13; DB 11; Length 343;
 Best Local Similarity 100.0%; Pred. No. 6.7e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 236 YSGKADVWSLGV 249
 |||||
 Db 229 YSGKADVWSLGV 242

RESULT 6

ID 015180 PRELIMINARY; PRT; 224 AA.
 AC 015180;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Phosphoprotein (Fragment).
 GN C8FW.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Thyroid;
 RA Wilkain F.; Suarez-Huerta N.; Robaye B.; Peetermans J.; Libert P.;
 RA "Characterization of a phosphoprotein whose mRNA is regulated by the
 RT mitogenic pathways in dog thyroid cells.";
 RL Eur. J. Biochem. 248:660-669(1997).
 DR EMBL; AJ000480; CA04119.1; -
 DR InterPro; IPR000719; Prot_kinase.
 DR pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 FT NON TER 1
 SQ SEQUENCE 224 AA; 25479 MW; 740CDD0905F86499B CRC64;

Query Match 3.6%; Score 13; DB 4; Length 224;
 Best Local Similarity 100.0%; Pred. No. 0.00043;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 237 YSGKADVWSLGV 249
 |||||
 Db 112 YSGKADVWSLGV 124

RESULT 7

ID 096R06 PRELIMINARY; PRT; 364 AA.
 AC 096R06;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE G-protein-coupled receptor induced protein GIG2 (Fragment).
 GN GIG2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; Tissue=Brain;
 RA Mayhew M.; von der Kammer H.; Klaudiny J.; Albrecht C.; Hoffmann B.;
 RA Nitich R.M.;
 RT "Identification of a novel nuclear factor GIG2, as an m1-acetylcholine
 RT receptor-induced gene.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF205438; AAG35664.1; -
 DR InterPro; IPR000719; Prot_kinase.
 DR pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Receptor; Transferase.
 FT NON TER 1
 SQ SEQUENCE 364 AA; 40377 MW; 8F9167FB76DFCD37 CRC64;

Query Match 3.6%; Score 13; DB 11; Length 364;
 Best Local Similarity 100.0%; Pred. No. 0.00066;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 237 YSGKADVWSLGV 249
 |||||
 Db 253 YSGKADVWSLGV 265

RESULT 8

ID 096R08 PRELIMINARY; PRT; 372 AA.
 AC 096R08;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE SKIPL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Kins-Tech E.; Wyllie D.H.; Qvarnstrom E.B.; Dover S.K.;
 RT "Identification of pro-inflammatory cytokine signalling network
 RT components by transcription expression screening.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF250310; AAK58174.1; -
 DR InterPro; IPR000719; Prot_kinase.
 DR pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 FT NON TER 1
 SQ SEQUENCE 372 AA; 40380 MW; 3E2B5C87A4F98FDB CRC64;

Query Match 3.6%; Score 13; DB 4; Length 372;
 Best Local Similarity 100.0%; Pred. No. 0.00067;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 237 YSGKADVWSLGV 249
 |||||

DB 260 YSGKADWWSLGV 272

RESULT 9

Q9H2Y8 PRELIMINARY; PRT; 372 AA.

AC Q9H2Y8; 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE G-protein-coupled receptor induced protein GIG2.

GN GIG2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBT_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RA Mayhau M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B., Nitsch R.M.;

RT "Identification of a novel nuclear factor GIG2, as an m1-acetylcholine receptor-induced gene."

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

RN EMBL; AF205437; AAG35663.1; -.

DR InterPro; IPR000719; Prot_Kinase.

DR Pfam; PF00069; Pkinase; 1.

DR ProDom; PD000001; Prot_Kinase; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KW ATP-binding; Receptor; Transferase

SW SEQUENCE 372 AA; 4108 MW; 5F54E50924B1365B CRC64;

Query Match 3.6%; Score 13; DB 4; Length 372; Best Local Similarity 100.0%; Pred. No. 0.0067; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 YSGKADWWSLGV 249

DB 260 YSGKADWWSLGV 272

RESULT 10

Q8K4K4 PRELIMINARY; PRT; 372 AA.

AC Q8K4K4; 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE TRB-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBT_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Kise-Toch E., Dempsey C., Jozsa V., Caunt J., Oxley K.M., Bagestiff S.M., Wylie D.H., Harte M., O'Neill L.A.J., Gwernstrom E.E., Dower S.K.;

RT "Mammalian homologs of Drosophila triblins (trb) control mitogen activated protein kinase signaling."

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

RN EMBL; AF358866; AAM45478.1; -.

DR InterPro; IPR000719; Prot_Kinase.

DR Pfam; PF00069; Pkinase; 2.

DR ProDom; PD000001; Prot_Kinase; 1.

DR SMART; SM00220; S_TKc.1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KW ATP-binding; Transferase.

SW SEQUENCE 372 AA; 41282 MW; 3A3DB82B46CD907F CRC64;

Query Match 3.6%; Score 13; DB 11; Length 372; Best Local Similarity 100.0%; Pred. No. 0.0067;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 YSGKADWWSLGV 249

DB 260 YSGKADWWSLGV 272

RESULT 11

Q91W04 PRELIMINARY; PRT; 372 AA.

AC Q91W04; 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Similar to phosphoprotein regulated by mitogenic pathways.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBT_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Breast tumor;

RA Strauberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=skin;

RX MEDLINE=22354683; PubMed=12466851;

RA The PANTOM Consortium.

RT the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR EMBL; BC006800; AA06800.1; -.

DR EMBL; AK028626; BAC26038.1; -.

DR InterPro; IPR000719; Prot_Kinase.

DR Pfam; PF00069; Pkinase; 1.

DR ProDom; PD000001; Prot_Kinase; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KW ATP-binding; Transferase.

SW SEQUENCE 372 AA; 41261 MW; AD29B8A640B462 CRC64;

Query Match 3.6%; Score 13; DB 11; Length 372; Best Local Similarity 100.0%; Pred. No. 0.0067; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 YSGKADWWSLGV 249

DB 260 YSGKADWWSLGV 272

RESULT 12

Q8H7S4 PRELIMINARY; PRT; 339 AA.

AC Q8H7S4; 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical protein.

GN OsJNB0081P02.7.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhardiidae; Oryzaceae; Oryza.

OX NCBT_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA Wang R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Sasaki C., Currie J., Collura K.;

RT "Rice Genomic Sequence."

RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC107226; AAN52742.1; -.

KW Hypothetical protein.

SO SEQUENCE 339 AA; 36383 MW; 4514842089AB607B CRC64;
 Query Match 2.8%; Score 10; DB 10; Length 339;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 AADVMSLGVA 250
 DB 235 AADVMSLGVA 244

RESULT 13

09F3C2 PRELIMINARY; PRT; 317 AA.

AC 09F3C2; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Purative oxidoreductase.
 GN SC07553 OR SCSP1.07.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.
 CX NCBI_TaxId=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Saunders D.C.; Harris D.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Cerdeno A.M.; Parkhill J.; Barrell B.G.; Rajandream M.A.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=97000351; PubMed=8643436;
 RX Redenbach M., Kiese H.M., Denepalte D., Eichner A., Cullum J.,
 RA Knaeht H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 Thompson N.R., James K.D., Harris D.E., Quail M.A., Kiese H.,
 Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 Huang C.-H., Kiese T., Lathe L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz B., Rajandream M.A., Rutherford K., Rutter S.,
 Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 DR EMBL; AL939132; CAC16433.1;
 DR InterPro; IPR002085; Adh_zn_family.
 DR Pfam; PF00107; adh_zinc_1.
 KW Complete proteome.
 SO SEQUENCE 317 AA; 31455 MW; 5AB9335A378352CD CRC64;

Query Match 2.5%; Score 9; DB 16; Length 317;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 LPAGISAA 288
 DB 107 LPAGISAA 115

RESULT 14

Q18041 PRELIMINARY; PRT; 562 AA.

ID Q18041
 AC Q18041; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical 63.4 Kda protein.
 GN C16B8.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Bentley D.;
 RT "The sequence of C. elegans cosmid C16B8.";
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U41031; AAA82618.2;
 DR HSP; P11362; 1FGK.
 DR WormPep; C16B8.1; CE27692.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR003306; WIF.
 DR Pfam; PF00069; Pkinase; 1.
 DR Pfam; PF02019; WIF; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00469; WIF; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 KW Hypothetical protein; ATP-binding; Kinase; Transferase;
 KW Tyrosine-protein kinase.
 SO SEQUENCE 562 AA; 63406 MW; F36B869F7F84C916 CRC64;

Query Match 2.5%; Score 9; DB 5; Length 562;
 Best Local Similarity 100.0%; Pred. No. 7.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 AADVMSLGVA 249
 DB 477 AADVMSLGVA 485

RESULT 15

Q9X266 PRELIMINARY; PRT; 583 AA.

ID Q9X266
 AC Q9X266; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Tyrosine kinase receptor-related protein Rtk precursor.
 GN Rtk.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_TaxId=6239;

```

KN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-B1etcl N2;
RX MEDLINE-99167506; PubMed-10066802;
RA Halford M.M., Oates A.C., Hibbs M.L., Wilks A.F., Stacker S.A.;
RT "Genomic structure and expression of the mouse growth factor receptor
   related to tyrosine kinases (RYK).";
RL J. Biol. Chem. 274:7379-7390(1999).
DR EMBL; AF133217; AAD24877.1; -.
DR HSBP; P11362; IFGK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR003306; WIF.
DR Pfam; PF00069; PKinase; 1.
DR Pfam; PF02019; WIF; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00469; WIF; 1_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Receptor; Signal; Transferase;
KW Tyrosine-protein kinase.
FT SIGNAL 1 18
FT CHAIN 19 583
FT SEQUENCE 583 AA, 65901 MW, F86DF8E47D24300 CRC64;
SQ

```

POTENTIAL. TYROSINE KINASE RECEPTOR-RELATED PROTEIN.

Query Match 2.5%; Score 9; DB 5; Length 583;
 Best Local Similarity 100.0%; Pred.No. 7.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 AADVMSLGIV 249
 |||||
 DB 498 AADVMSLGIV 506

Search completed: January 15, 2004, 15:06:48
 Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:01:57 | Search time 17 Seconds

(without alignment)
990.327 Million cell updates/sec

Title: US-09-909-474d-2

Sequence: 1 MRATPLAPAGSLSRKKRLG.....GLGLDEARBSRGREVLVYG 358

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	274	76.5	358	1 NIPK_HUMAN	Q96RU7 homo sapien
2	25	7.0	349	1 NIPK_RAT	Q9WQ66 rattus norv
3	22	6.1	354	1 NIPK_MOUSE	Q8K4K2 mus musculu
4	8	2.2	174	1 HUNB_DROMU	Q46250 drosophila
5	8	2.2	376	1 P46G_HUMAN	Q9BY94 homo sapien
6	8	2.2	431	1 NRKA_TRYBB	Q08942 trypanosoma
7	8	2.2	431	1 NRKB_TRYBB	Q03428 trypanosoma
8	7	2.0	82	1 YKPA_YEAST	Q60200 saccharomyc
9	7	2.0	91	1 B310_ADE03	P11318 human adeno
10	7	2.0	91	1 B310_ADE03	P11314 human adeno
11	7	2.0	116	1 MBRT_PSEAE	P04140 pseudomonas
12	7	2.0	156	1 ATPF_VIBAL	P12988 vibrio algi
13	7	2.0	195	1 TATB_XANCP	Q9P3H5 xanthomonas
14	7	2.0	201	1 CTPI_HUMAN	Q16613 homo sapien
15	7	2.0	204	1 PAAD_BACSU	P44404 bacillus su
16	7	2.0	209	1 VOIT_ECOLI	P76657 escherichia
17	7	2.0	219	1 R819_AQUAB	Q66435 aquifex aeo
18	7	2.0	246	1 ADC_CLOBR	Q9P4K1 clostridium
19	7	2.0	295	1 NMCR_BNTCL	P52676 enterobacte
20	7	2.0	315	1 OMPF_ECOLI	P23676 escherichia
21	7	2.0	316	1 L767_CASBL	Q09517 caenorhabdi
22	7	2.0	324	1 HEM3_PAEWA	Q09110 pseudocilli
23	7	2.0	332	1 LPXK_PSEAB	Q9H2M3 pseudomonas
24	7	2.0	344	1 RLAD_ABRPB	Q9Y968 aeropyrum p
25	7	2.0	349	1 NRUB_KLRPO	P10045 klebsiella
26	7	2.0	370	1 M10W_AFLCA	P15513 a myomodul
27	7	2.0	378	1 LUXE_VIRHA	P14286 vibrio harv
28	7	2.0	397	1 DXR_HABIN	P44055 haemophilus
29	7	2.0	419	1 CPBY_SYNPY	Q02174 haemophilus
30	7	2.0	419	1 MPK7_HUMAN	Q14733 homo sapien
31	7	2.0	435	1 FUZ7_USRMA	Q99078 usciiliago ma
32	7	2.0	441	1 NFIX_MESAU	P13623 mesocricetu
33	7	2.0	451	1 CBPS_STRGR	P18143 streptomyces

34	7	2.0	460	1 V4TB_METAC	Q8L1J0 methanosaarc
35	7	2.0	460	1 V4TB_METBA	P22663 methanosaarc
36	7	2.0	460	1 V4TB_METMA	Q60187 methanosaarc
37	7	2.0	487	1 HBP_DROME	Q23977 drosophila
38	7	2.0	507	1 GTR6_HUMAN	Q9UGQ3 homo sapien
39	7	2.0	515	1 AMPA_MYCTU	Q10401 mycobacteri
40	7	2.0	519	1 GIG2_SOLTU	P55242 solanum tub
41	7	2.0	535	1 TX21_HUMAN	Q9U117 homo sapien
42	7	2.0	545	1 MUTL_THETH	Q9T454 thermus the
43	7	2.0	595	1 STD_BACID	Q9K491 bacillus ha
44	7	2.0	605	1 WIS1_SCHPO	P31886 eschizosacch
45	7	2.0	614	1 MGM_EUBBA	Q59268 eubacterium

ALIGNMENTS

RESULT 1
ID NIPK_HUMAN STANDARD, PRT: 358 AA.
AC Q96RU7, Q9TAT6, Q9H5M8, Q9NUD2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neuronal cell death inducible putative kinase (SKIP3).
GN NIPK OR C200RP97.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiebaaba A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Oktani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kiebaaba A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Oktani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley C.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carter C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levanon S.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McWhirry A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prichard S.R., Plumb R.W., Ramsey H.,
RA Rice C.M., Rose M.T., Scott C.E., Sghra H.K., Showlken R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulten J.B.,
RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Waller D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).

```

RN [4]
RP SEQUENCE FROM N.A.
RC TIS8U8-Cervix, and Musclic;
RA MEDLINE=22388257; PubMed=12477932;
RA Schreiber R.L., Feinberg A.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.P., Bhat N.K.,
RA Hopline R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Patney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalhus D.B.,
RA Schermer A., Schein J.E., Jones S.J.W., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: May play an important role in a common pathway leading
CC to programmed neuronal cell death. Does not appear to function in
CC the programmed death of non-neuronal cells. May serve as an
CC endogenous antagonist competing for substrate with functional
CC kinases that act to promote neuronal cell survival (by
CC similarity).
CC -1- SIMILARITY: Contains 1 protein kinase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF250311; AAK58175.1; -
DR EMBL; AK026945; BAB15597.1; -
DR EMBL; AL034548; CAB15634.1; -
DR EMBL; BC019363; AAH19363.1; -
DR EMBL; BC027484; AAH27484.1; -
DR Genew; HGNC:16228; C20orf97;
DR InterPro; IPR000719; Prot_Kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM Apoptosis.
FT DOMAIN 68 316 PROTEIN KINASE.
FT CONFLICT 84 84 Q -> R (IN REF. 4; AAH27484).
FT CONFLICT 105 105 L -> P (IN REF. 2).
FT CONFLICT 114 114 L -> V (IN REF. 1).
FT CONFLICT 194 195 ER -> DREK (IN REF. 1).
SO SEQUENCE 358 AA; 39577 MW; CRISF89A81BD63 CRC64;

Query Match 76.5%; Score 274; DB 1; Length 358;
Best Local Similarity 100.0%; Pred. No. 8.2e-272;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 ALHCPGTGTYCKYPPQVQALAVLEPYARLPKHGVNAPTEVLAQTOLLVAFPTRTGDM 144
DB 85 ALHCPGTGTYCKYPPQVQALAVLEPYARLPKHGVNAPTEVLAQTOLLVAFPTRTGDM 144
QY 145 HSLVRSRRIRIPEPEAAVLFROMATLAHQHGLVLRDLKLCRFVPADEBKKQVLENTLE 204
DB 145 HSLVRSRRIRIPEPEAAVLFROMATLAHQHGLVLRDLKLCRFVPADEBKKQVLENTLE 204
QY 205 DSCVLTGPDSDLMDKACPAVYGPBILSSRSYSGKADVWSLGVALTFTMLAGHYPPQDS 264
DB 205 DSCVLTGPDSDLMDKACPAVYGPBILSSRSYSGKADVWSLGVALTFTMLAGHYPPQDS 264

```

```

DB 205 DSCVLTGPDSDLMDKACPAVYGPBILSSRSYSGKADVWSLGVALTFTMLAGHYPPQDS 264
QY 265 BPVLLFGKIRRGAYALPAGLSAPARCLVRCILRRBPARRLTATYGLLHPWLRODMPPLAP 324
DB 265 BPVLLFGKIRRGAYALPAGLSAPARCLVRCILRRBPARRLTATYGLLHPWLRODMPPLAP 324
QY 325 TRSHHMAAOVPPDGLGLDEARREBEGDREVVLYG 358
DB 325 TRSHHMAAOVPPDGLGLDEARREBEGDREVVLYG 358

RESULT 2
NIPK_RAT
ID NIPK_RAT STANDARD; PRT; 349 AA.
AC 09W06;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuronal cell death inducible putative kinase.
GN NIPK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TIS8U8-Neuron;
RA MEDLINE=99262087; PubMed=10329375;
RA Mayumi-Matsuda K., Kojima S., Suzuki H., Sakata T.;
RT "Identification of a novel kinase-like gene induced during neuronal
RT cell death."
RL Biochem. Biophys. Res. Commun. 258:260-264(1999).
CC -1- FUNCTION: May play an important role in a common pathway leading
CC to programmed neuronal cell death. Does not appear to function
CC in the programmed death of non-neuronal cells. May serve as an
CC endogenous antagonist competing for substrate with functional
CC kinases that act to promote neuronal cell survival.
CC -1- TISSUE SPECIFICITY: Detected only in the lung, kidney and testis.
CC heart, brain, spleen, liver, skeletal muscle, kidney and testis.
CC -1- INDUCTION: Expression induced during programmed cell death evoked
CC in neuronal cells by NGF-depletion.
CC -1- SIMILARITY: Contains 1 protein kinase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB020967; BAA7582.1; -
DR InterPro; IPR000719; Prot_Kinase.
DR Pfam; PF00069; pkinase; 2.
DR ProDom; PD000001; Prot_Kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM Apoptosis.
FT DOMAIN 63 310 PROTEIN KINASE.
SO SEQUENCE 349 AA; 38602 MW; 3050F9BF946D815 CRC64;

Query Match 7.0%; Score 25; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.9e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 TGPDDSLMDKACPAVYGPBILSSR 234
DB 210 TGPDDSLMDKACPAVYGPBILSSR 229

RESULT 3
NIPK_MOUSE
ID NIPK_MOUSE STANDARD; PRT; 354 AA.
AC 08K4K2; Q921E7;

```

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Neuronal cell death inducible putative Kinase (TRB-3).
 GN NIPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Makedo I., Otsu N., Saito R., Suzuki H., Yamana K., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirrl L.M., Knapin A., Matsuda H., Batalov S., Betsel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbali L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazier K.S.,
 RA Gaasterland T., Gariboldi M., Glast C., Godzik A., Gough J.,
 RA Grimmond S., Guentrich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Keddie J., Kersey R.M., King B.L.,
 RA Konegaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,
 RA Maglott D.R., Maltaise L., Marchionni L., McKenzie L., Miki H.,
 RA Nagelsmith T., Numata K., Okido T., Pavan M.J., Pereira G., Pesole G.,
 RA Percevaevy N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner R., Whitlested C., Wang Y., Watanabe Y., Wells C.,
 RA Wilmfing L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carlinici P., Hayatsu N.,
 RA Hirozono-Kishikawa T., Kono H., Nakanura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawaj J., Aizawa K., Aizawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Seebak D., Shibata K., Shingawa K.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Gerbautsberg R.L., Feilgenfeld E.A., Gronow L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altshuler S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Hsieh N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.R., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Caavaant T.L., Schetz T.S.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carlinici P., Prange C.,
 RA Roha S.G., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Baskin S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richardson S., Worley K.C., Hale S.G., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalón D.K., Muzny D.M., Sodegren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Holton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodierfeld A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.B.,
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May play an important role in a common pathway leading

```

CC      to programmed neuronal cell death. Does not appear to function in
CC      the programmed death of non-neuronal cells. May serve as an
CC      endogenous antagonist competing for substrate with functional
CC      kinases that act to promote neuronal cell survival (By
CC      similarity).
CC      -1- SIMILARITY: Contains 1 protein kinase domain.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcement/1986-81b.ch).
CC      or send an email to license@ebi.ac.uk.
CC
CC      EMBL, AF358668; AAA45476.1; -.
CC      EMBL, AK089931; BAC41002.1; -.
CC      EMBL, BC012955; AAA12955.1; -.
CC      InterPro: IPR000719; Prot_kinase.
CC      InterPro: IPR002290; Ser_thr_kinase.
CC      Pfam: PF00069; Pkinase; 2.
CC      ProDom: PD000001; Prot_kinase; 1.
CC      SMART: SM00220; S_TKc; 1.
CC      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC      Apoptosis.
CC      KW      DOMAIN.
CC      FT      68           315      PROTEIN KINASE.
CC      FT      CONFLICT 157       157      S -> P (IN REF. 3) .
CC      FT      CONFLICT 219       219      K -> T (IN REF. 1) .
CC      FT      CONFLICT 239       264      MISSING (IN REF. 3) .
CC      FT      CONFLICT 301       354      SERVALGILHPMLRHDHGSPPOSDREMDQVDPDQ
CC      FT      FT      LEERAESESVGLYG -> CRATCGGNPLASIVERSRPSLS
CC      STV (IN REF. 2) .
CC      SQ      SEQUENCE 354 AA; 39022 MW; 2CB283FC119F859F CRC64;
CC
CC      Query Match 6.1%; Score 22; DB 1; Length 354;
CC      Best Local Similarity 100.0%; Pred. No. 2,2e-14;
CC      Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC      QY      213 DDSLMDRGRACPAAYVGPRLSSR 234
CC      Db      213 DDSLMDRGRACPAAYVGPRLSSR 234
CC
CC      RESULT 4
CC      HUNB_DROMU STANDARD; PRT; 174 AA.
CC      ID      HUNB_DROMU
CC      AC      046250; 046251;
CC      DT      16-OCT-2001 (Rel. 40, Created)
CC      DT      16-OCT-2001 (Rel. 40, Last sequence update)
CC      DT      15-SEP-2003 (Rel. 42, Last annotation update)
CC      DE      Hunchback protein (Fragments) .
CC      GN      HB.
CC      OS      Drosophila mulleri (Fruit fly) .
CC      OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC      OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC      OC      Ephydroidea; Drosophilidae; Drosophila.
CC      OX      NCBI_TaxID=7231;
CC      RN      [1]
CC      RP      SEQUENCE FROM N.A.
CC      RA      Baker R.H., Desalle R.;
CC      RT      "Multiple sources of character information and the phylogeny of
CC      RT      Hawaiian Drosophilids.";
CC      RL      Syst. Biol. 46:654-673(1997).
CC      CC      -1- FUNCTION: GAP CLASS SEGMENTATION PROTEIN THAT CONTROLS DEVELOPMENT
CC      CC      OF HEAD STRUCTURES (BY SIMILARITY) .
CC      CC      -1- SUBCELLULAR LOCATION: Nuclear (By similarity) .
CC      CC      -1- SIMILARITY: BELONGS TO THE HUNCHBACK FAMILY OF C2H2-TYPE ZINC-
CC      CC      FINGER PROTEINS.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U93014; AAC03262.1; -
 DR EMBL; U93015; AAC03263.1; -
 DR P1YBases; FBgn00233760; Dmnl\hb.
 KW Developmental protein; Gap protein; Zinc-finger;
 KM Metal-binding; DNA-binding; Repeat; Nuclear protein.
 PT NON_TER 1
 FT DOMAIN 1
 FT NON_CONS 58 67 POLY-GLN.
 FT NON_TER 93 94
 FT NON_TER 174 174
 SQ SEQUENCE 174 AA; 19265 MW; DA27CF6CC8CC8368 CRC64;

Query Match 2.2%; Score 8; DB 1; Length 174;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 RATVAVTA 65
 DB 118 RATVAVTA 125

RESULT 5
 P6G_HUMAN STANDARD; PRT; 376 AA.
 AC Q9BYG4;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DE 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Partitioning defective-6 homolog gamma (PAR-6 gamma) (PAR6D).
 GN PAR6G OR PAR6G.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND INTERACTION WITH RAC1; CDC42; PRKCI AND PRK2.
 RC TISSUE=Neuroblastoma;
 RX PubMed=11260256;
 RA Noda Y., Takeya R., Ohno S., Naito S., Ito T., Sumimoto H.;
 RT "Human homologues of the Caenorhabditis elegans cell polarity protein
 PAR6 as an adaptor that links the small GTPases Rac and Cdc42 to
 RT atypical protein kinase C.";
 RL Genes Cells 6:107-119(2001).
 RN [2]
 RP INTERACTION WITH ARHO.
 RA MEDLINE=20394296; PubMed=10934474;
 RX Joberty G., Petersen C., Gao L., Macara I.G.;
 RT "The cell-polarity protein Par6 links Par3 and atypical protein kinase
 C to Cdc42.";
 RL Nat. Cell Biol. 2:531-539(2000).
 CC -1- FUNCTION: Adapter protein involved in asymmetrical cell division
 CC and cell polarization processes. May play a role in the formation
 CC of epithelial tight junctions. The PAR6-PAR3 complex links GTP-
 CC bound Rho small GTPases to atypical protein kinase C proteins (By
 CC similarity).
 CC -1- SUBUNIT: Interacts with PAR63 (Probable). Interacts with GTP-bound
 CC forms of CDC42, ARHO/TC10 and RAC1. Interacts with the N-terminal
 CC part of PRKCI and PRK2.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: Widely expressed, with a higher expression in
 CC fetal and adult kidney.
 CC -1- DOMAIN: The pseudo-CRIB domain together with the PDZ domain is
 CC required for the interaction with Rho small GTPases (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the PAR6 family.
 CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
 CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB044556; BAB40757.1; -
 DR Genew; HGNC:16076; PAR6G.
 DR InterPro; IPR001478; PDZ.
 DR Pfam; PF00595; PDZ; 1.
 DR SMART; SM00666; PB1; 1.
 DR SMART; SM00228; PDZ; 1.
 DR PROSITE; PS50106; PDZ; 1.
 KW Cell cycle; Cell division; Tight junction; Membrane.
 KM DOMAIN 134 151
 FT DOMAIN 158 251
 FT DOMAIN 127 254
 SQ SEQUENCE 376 AA; 40883 MW; B5A4BF521403CA11 CRC64;
 (BY SIMILARITY)

Query Match 2.2%; Score 8; DB 1; Length 376;
 Best Local Similarity 100.0%; Pred. No. 4.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APAGSLSR 15
 DB 318 APAGSLSR 325

RESULT 6
 NRKA_TRYB STANDARD; PRT; 431 AA.
 ID NRKA_TRYB
 AC Q08942;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative serine/threonine-protein kinase A (EC 2.7.1.37).
 GN NRKA.
 OS Trypanosoma brucei brucei.
 OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxId=5702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SATRO 164 / Isolate ISTARI;
 RX MEDLINE=93295429; PubMed=8515773;
 RA Gale M.J. Jr., Parsons M.;
 RT "A Trypanosoma brucei gene family encoding protein kinases with
 RT catalytic domains structurally related to Nekk1 and NIMA.";
 RL Mol. Biochem. Parasitol. 59:111-122(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L03778; AAB59252.1; -
 DR PIR; T11854; T11854.
 DR HSSP; O63450; IA06.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR SMART; SM00233; PH; 1.


```

DR SMART; SM00220; S TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 20 279
FT NP_BIND 331 429 PH. (BY SIMILARITY).
FT BINDING 26 34 ATP (BY SIMILARITY).
FT ACT_SITE 147 147 ATP (BY SIMILARITY).
FT VARIANT 192 192 A -> P (IN STRAIN TREU66).
FT VARIANT 199 199 K -> L (IN STRAIN TREU66).
SQ SEQUENCE 431 AA; 47915 MW; 2868B70275864224 CRC64;

Query Match 2.2%; Score 8; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ADVMSLGV 249
DB 205 ADVMSLGV 212

RESULT 7
NRKB TRYBB
ID NRKB TRYBB STANDARD; PRT; 431 AA.
AC Q03428;
DT 01-CT-1994 (Rel. 30; Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Pucative serine/threonine-protein kinase B (EC 2.7.1.37).
GN NRKB.
OS Trypanosoma brucei brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OC NCBI_TaxId=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=isoate TREU66;
RX MEDLINE=93295429; PubMed=8515773;
RA Gale M.J. Jr.; Parsons M.;
RT "A Trypanosoma brucei gene family encoding protein kinases with
RT catalytic domain structurally related to Nekl and NIMA.";
RL Mol. Biochem. Parasitol. 59:111-122(1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; L03777; AAB59253.1; -
DR InterPro; IPR01849; PH.
DR InterPro; IPR00719; Prot_Kinase.
DR InterPro; IPR002290; Ser_Thr_Kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; PKinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_Kinase; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 20 279
FT NP_BIND 331 429 PH. (BY SIMILARITY).
FT BINDING 26 34 ATP (BY SIMILARITY).
FT ACT_SITE 147 147 ATP (BY SIMILARITY).
FT VARIANT 192 192 A -> P (IN STRAIN TREU66).
FT VARIANT 199 199 K -> L (IN STRAIN TREU66).
SQ SEQUENCE 431 AA; 47915 MW; 2868B70275864224 CRC64;

```

```

FT DOMAIN 331 429 PH. (BY SIMILARITY).
FT NP_BIND 26 34 ATP (BY SIMILARITY).
FT BINDING 49 49 ATP (BY SIMILARITY).
FT ACT_SITE 147 147 ATP (BY SIMILARITY).
SQ SEQUENCE 431 AA; 48172 MW; F6B4B3BF959PF74 CRC64;

Query Match 2.2%; Score 8; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ADVMSLGV 249
DB 205 ADVMSLGV 212

RESULT 8
YKFA YEAST
ID YKFA YEAST STANDARD; PRT; 82 AA.
AC O60200;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Hypothetical 9.3 kDa protein in ORF1-ORF4 intergenic region.
GN YKL053C-A.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94378723; PubMed=8091862;
RA Raemussen S.W.;
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the
RT FBAl and TOA2 genes, an open reading frame (ORF) similar to a
RT translationally controlled tumour protein, one ORF containing motifs
RT also found in plant storage proteins and 13 ORFs with weak or no
RT homology to known proteins.";
RL Yeast 10:563-568(1994).
CC -1- SIMILARITY: BELONGS TO THE UPF0203 (15B1.1) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Z28054; CAA81891.1; ALT_INIT.
DR EMBL; Z28052; CAA81889.1; ALT_INIT.
DR SGD; S0007243; YKL053C-A.
DR GO; GO:0007005; P:mitochondrion organization and biogenesis; IMP.
DR Pfam; PF03254; UPF0203; 1.
KW Hypothetical protein.
SQ SEQUENCE 82 AA; 9296 MW; 6D8EA7CDD89FB73E CRC64;

Query Match 2.0%; Score 7; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 LDEARB 348
DB 62 LDEARB 68

RESULT 9
B310 ADE03
ID B310 ADE03 STANDARD; PRT; 91 AA.
AC P1318;
DT 01-JUL-1989 (Rel. 11; Created)
DT 01-JUL-1989 (Rel. 11; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)

```

```

DE Early E3B 10.4 kDa protein precursor.
OS Human adenovirus type 3.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_Taxid=45659;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87219876; PubMed=3582978;
RA Signaes C., Anusjaervi G., Pettersson U.;
RT "Region E3 of human adenoviruses; differences between the oncogenic
RL Gene 50:173-184(1986).
RN [2]
RP IDENTIFICATION OF PROTEIN.
RX MEDLINE=90112650; PubMed=2296083;
RA Tollefson A.E., Krajsci P., Yet S., Carlin C.R., Wold W.S.M.;
RT "A 10,400-molecular-weight membrane protein is coded by region E3 of
RL adenovirus."
RN [3]
RP J. Virol. 64:794-801(1990).
CC -1- FUNCTION: DOWN-REGULATES THE EGF RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
CC reticulum (Probable).
CC -1- SIMILARITY: TO EQUIVALENT PROTEIN IN OTHER ADENOVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@ib-sib.ch).
CC -----
DR EMBL, M15952; AAA42487.1; -.
DR InterPro: IPR005041; Adeno_E3B.
DR Pfam: PF03376; Adeno_E3B; 1.
DR ProDom: PD006617; Adeno_E3B; 1.
DR Early protein, Transmembrane; Signal; Endoplasmic reticulum.
FT SIGNAL 1 22
FT CHAIN 23 91
FT DOMAIN 23 34
FT TRANSMEM 35 60
FT DOMAIN 61 91
FT DOMAIN 91 91
SQ SEQUENCE 91 AA; 10356 MW; 10F5281008DB2C41 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 TAVATAS 66
DB 22 TAVATAS 28

RESULT 10
E310_ADE07 STANDARD; PRT; 91 AA.
AC P15134;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Early E3B 10.4 kDa protein precursor.
OS Human adenovirus type 7.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_Taxid=10519;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Gomen;
RX MEDLINE=89073758; PubMed=2649239;
RA Hong J.S., Mullis K.G., Engler J.A.;
RT "Characterization of the early region 3 and fiber genes of Ad7."
RN [2]
RP Virology 167:545-553(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=87-922;

```

```

RA Kajian A.B., Madell G.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DOWN-REGULATES THE EGF RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
CC reticulum (Probable).
CC -1- SIMILARITY: TO EQUIVALENT PROTEIN IN OTHER ADENOVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@ib-sib.ch).
CC -----
DR EMBL, M23696; AAA53251.1; -.
DR EMBL, Z48954; CAA88813.1; -.
DR PIR, G31830; ERAD27.
DR PIR, G31830; ERAD27.
DR InterPro: IPR005041; Adeno_E3B.
DR Pfam, PF03376; Adeno_E3B; 1.
DR ProDom, PD006617; Adeno_E3B; 1.
DR Early protein, Transmembrane; Signal; Endoplasmic reticulum.
FT SIGNAL 1 22
FT CHAIN 23 91
FT DOMAIN 23 34
FT TRANSMEM 35 60
FT DOMAIN 61 91
FT DOMAIN 91 91
SQ SEQUENCE 91 AA; 10330 MW; 0C8534608CE2C54 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 TAVATAS 66
DB 22 TAVATAS 28

RESULT 11
MERT_PSEAE STANDARD; PRT; 116 AA.
ID MERT_PSEAE
AC P04140;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Mercutic transport protein (Mercutic ion transport protein).
GN MERT.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_Taxid=287;
RN [1]
RP SEQUENCE FROM N.A.
RX TRANSPOSON=Tn501;
RX MEDLINE=85014891; PubMed=6091128;
RA Miara T.K., Brown N.L., Fritzing D.C., Fridmore R.D., Barnes W.M.,
RA Haberer L., Silver S.;
RT "Mercutic ion-resistance operons of plasmid R100 and transposon
RT Tn501: the beginning of the operon including the regulatory region
RT and the first two structural genes."
RL Proc. Natl. Acad. Sci. U.S.A. 81:5975-5979(1984).
CC -1- FUNCTION: INVOLVED IN MERCUTIC TRANSPORT. PASSES A HG(2+) ION
CC FROM THE PERIPLASMIC MERT PROTEIN TO THE MERCUTIC REDUCTASE
CC (MERA).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/> or send an email to license@lsb-sib.ch).

CC -----

DR EMBL, Z00027, CAAT7321.1, -

DR EMBL, K02503, AAA27433.1, -

DR PIR, A04457, QOPSHT.

DR InterPro, IPR003457, Transprt_MerT.

DR Pfam, PF02411, MerT.1.

KM Transport, Transposable element; Mercuric resistance; Inner membrane;

KM Mercury; Plasmid; Transmembrane.

FT TRANSMEM 16 36 POTENTIAL.

FT TRANSMEM 44 64 POTENTIAL.

FT TRANSMEM 94 114 POTENTIAL.

FT METAL 24 24 HG(2+) (POTENTIAL).

FT METAL 25 25 HG(2+) (POTENTIAL).

FT METAL 76 76 HG(2+) (POTENTIAL).

FT METAL 82 82 HG(2+) (POTENTIAL).

SO SEQUENCE 116 AA; 12498 MW; 905B5492AFB086C CRC64;

Query Match 2.0%; Score 7; DB 1; Length 116;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 LAVALPEY 111

Db 44 LAVALPEY 50

RESULT 12

ID ATYP VIBAL STANDARD; PRT; 156 AA.

AC P12989;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE ATP synthase B chain (EC 3.6.3.14).

GN ATP OR UNCF.

OS *Vibrio alginolyticus*.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrrio.

OX NCBI_TaxId=663;

OK

RP SEQUENCE FROM N.A.

RC STRAIN=138-2;

RX MEDLINE=90016889; PubMed=2529481;

RA Krumholz L.R., Baer U., Simoni R.D.;

RT "Nucleotide sequence of the unc operon of *Vibrio alginolyticus*."

CC Nucleic Acids Res. 17:7993-7994(1989).

CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate + H(+) (out).

CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0) HAS THREE MAIN SUBUNITS: A, B AND C.

CC -1- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.lsb-sib.ch/announce/> or send an email to license@lsb-sib.ch).

CC -----

DR EMBL, X16050, CAA34177.1, -

DR PIR, S06078, S06078.

DR HSRP, P00859, 189U.

DR InterPro, IPR005864, ATPaseB.

DR InterPro, IPR002146, ATPaseB_B/B_sub.

DR Pfam, PF00430, ATP-synt_B_1.

DR TIGRfam, TIGR01144, ATP_synt_b_1.

KM Hydrogen ion transport; Transmembrane; CF(0).

FT TRANSMEM 2 22 POTENTIAL.

SO SEQUENCE 156 AA; 17532 MW; 89057460643D8CF1 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 156;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 342 LDEARER 348

Db 87 LDEARER 93

RESULT 13

ID TATB_XANCP STANDARD; PRT; 195 AA.

AC Q8P3H9;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Sec-independent protein translocase protein tatb homolog.

GN TATB OR XC04092.

OS *Xanthomonas campestris* (pv. *campestris*).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xanthomonas.

OX NCBI_TaxId=340;

OK

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33913 / NCPPB 528;

RX MEDLINE=22022145; PubMed=12024217;

RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R., Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F., Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A., Camarotte G., Canavan F., Cardoso J., Chabergo F., Clippin L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Doroty H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A., Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F., Locali R.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M., Martins E.C., Medeiros J., Menck C.F.M., Miyaki C.Y., Moon D.H., Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R., Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F., Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira B.C., Tezza R.I.D., RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F., RA Setubal J.C., Kitajima J.P.;

RT "Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities."

CC Nature 417:459-463(2002).

CC -1- FUNCTION: Required for correct localization of precursor proteins bearing signal peptides with the twin arginine conserved motif S/T-R-R-X-F-L-R. This sec-independent pathway is termed TAR for twin-arginine translocation system. This system mainly transports proteins with bound cofactors that require folding prior to export (by similarity).

CC -1- SUBCELLULAR LOCATION: Inner-membrane bound (Probable).

CC -1- SIMILARITY: Belongs to the tatb family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.lsb-sib.ch/announce/> or send an email to license@lsb-sib.ch).

CC -----

DR EMBL, AB012532, AA043313.1, -

DR HAMAP, MP_00237, -1.

DR InterPro, IPR003369, MtaA_Hcfl06.

DR InterPro, IPR003998, TatB.

DR Pfam, PF02416, MtaA_Hcfl06, 1.

DR PRINTS, PR01506, TATBPROTEIN.

DR TIGRfam, TIGR01410, tatb; 1.

KM Transport; Protein transport; Translocation; Transmembrane;

KM Inner membrane; Complete proteome.

FT TRANSMEM 1 21 POTENTIAL.

SQ SEQUENCE 195 AA; 20929 MW; 8DDDB6F526AD390 CRC64;
 Query Match 2.0%; Score 7; DB 1; Length 195;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ATP1AAP 9
 DB 112 ATP1AAP 118

RESULT 14
 ID CTF1_HUMAN STANDARD; PRT; 201 AA.
 AC Q16619;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cardiotrophin-1 (CT-1).
 GN CTF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=96429882; PubMed=8833032;
 RA Penica D., Swanson T.A., Shaw K.J., Kuang W.-J., Gray C.L.,
 RA Beatty B.G., Wood W.I.;
 RT "Human cardiotrophin-1: protein and gene structure, biological and
 RT binding activities, and chromosomal localization.";
 RL Cytokine 8:183-189(1996).
 CC -1 FUNCTION: INDUCES CARDIAC MYOCYTE HYPERTROPHY IN VITRO. BINDS TO
 CC AND ACTIVATES THE LEUKEMIA INHIBITORY FACTOR RECEPTOR (LIF
 CC RECEPTOR/GP 130 RECEPTOR COMPLEX).
 CC -1 SUBCELLULAR LOCATION: Secreted.
 CC -1 TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, SKELETAL MUSCLE,
 CC PROSTATE AND OVARY. LOWER LEVELS IN LUNG, KIDNEY, PANCREAS,
 CC THYMUS, TESTIS AND SMALL INTESTINE. LITTLE OR NO EXPRESSION IN
 CC BRAIN, PLACENTA, LIVER, SPLEEN, COLON OR PERIPHERAL BLOOD
 CC LYMPHOCYTES.
 CC -1 SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC DR EMBL; U43033; AAD12173.1; -;
 CC DR EMBL; U43031; AAD12173.1; JOINED.
 CC DR EMBL; U43032; AAD12173.1; JOINED.
 CC DR EMBL; U43030; AAB55229.1; -;
 CC DR PIR; G02312; G02312.
 CC DR Genew; HGNC:2499; CTF1.
 CC DR MIM; 600435; -;
 CC DR GO; GO:0005576; C:extracellular; TAS.
 CC DR GO; GO:0005146; F:leukemia inhibitory factor receptor ligand . . .; TAS.
 CC DR GO; GO:0008283; P:cell proliferation; TAS.
 CC DR GO; GO:0007267; P:cell-cell signaling; TAS.
 CC DR GO; GO:0007517; P:muscle development; TAS.
 CC DR GO; GO:0007399; P:neurogenesis; TAS.
 CC DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
 CC KW Cytokine; Polymorphism.
 CC FT VARIANT 92 92 A -> T (IN dbSNP:2234933).
 CC FT /FTID=VAR_014938.
 CC SQ SEQUENCE 201 AA; 21227 MW; 0235AB5745F675F CRC64;
 Best Local Similarity 100.0%; Pred. No. 28;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 AGLSAPA 288
 DB 69 AGLSAPA 75

RESULT 15
 ID PAD_BACSU STANDARD; PRT; 204 AA.
 AC P94404;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable aromatic acid decarboxylase (EC 4.1.1.-).
 GN YCB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124189; PubMed=8969502;
 RA Yamane K., Kumano M., Kurita K.;
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denton F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson F.T.,
 RA Ertlan K.D., Ertling J., Fabre C., Ferrari B., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaeser P., Goffeau A., Golligly E.J., Grand G.,
 RA Giuseppe H., Guy B.J., Haga K., Haele J., Harwood C.R., Hensel A.,
 RA Hilbert G., Holst S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kashiwara Y., Kleier-Bianchini M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesl D., Nakel S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Priesen B., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan B., Schleich S., Schroeder R., Scofield F.,
 RA Sekiguchi J., Sekowka A., Seror S.J., Serrif P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi B., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Taperia P., Tognoni A.,
 RA Tobias V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambut R., Wedler B., Wedler H., Wellzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasuno K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein H., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -1 SIMILARITY: BELONGS TO THE POLYPYRROL P-HYDROXYBENZONATE /
 CC PHENYLACETIC ACID DECARBOXYLASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----

CC -----
DR EMBL/ D50453; BAA08996.1; -
DR EMBL/ Z99105; CAB12157.1; -
DR EMBL/ Z99106; CAB12171.1; -
DR PIR/ G69761; G69761.
DR Subtilase; BG12023; YC1B.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; Ubix.
DR Pfam/ PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; Ubix; 1.
KM Hypothetical protein; Lyase; Decarboxylase; Complete proteome.
SQ SEQUENCE 204 AA; 22539 MW; F6C1A5219365D728 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 ERKLV 200
|||
Db 124 ERKLV 130

Search completed: January 15, 2004, 15:05:20
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:02:32 ; Search time 21 Seconds

(without alignments)
1639.447 Million cell updates/sec

Title: US-09-909-474D-2

Perfect score: 358

Sequence: 1 MRAITPLAAPAGSLSRKKRLR.....GLGLDREAREEGCDREVLYG 358

Scoring table: OLIGO

Searched: 283308 seqs, 9616682 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Filtering filter 45 summaries

Database: PIR 76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.5	548	2 T15523	hypothetical prote
2	8	2.2	190	2 A95405	conserved hypothet
3	8	2.2	431	2 T11854	protein kinase (EC
4	8	2.2	725	2 A83266	conserved hypothet
5	8	2.2	813	2 A33258	cation-transportin
6	7	2.0	33	2 T36211	probable excisiona
7	7	2.0	66	2 H44057	early E3 10.3K pro
8	7	2.0	66	2 T36199	probable excisiona
9	7	2.0	72	2 AD3056	hypothetical prote
10	7	2.0	80	2 H89990	hypothetical prote
11	7	2.0	86	2 S78725	protein YKL053c-a
12	7	2.0	91	1 ERAD27	early E3 10.3K pro
13	7	2.0	91	1 ERAD77	early E3 10.3K pro
14	7	2.0	116	1 QP8HT	hypothetical prote
15	7	2.0	123	2 B69418	hypothetical prote
16	7	2.0	131	2 E72649	hypothetical prote
17	7	2.0	134	2 D84672	hypothetical prote
18	7	2.0	143	2 G95179	conserved hypothet
19	7	2.0	154	2 AH3143	hypothetical prote
20	7	2.0	156	1 S06078	H+-transporting tw
21	7	2.0	158	1 D98144	conserved hypothet
22	7	2.0	162	2 G98046	ribosomal protein
23	7	2.0	166	2 H70300	cardiotrophin-1 -
24	7	2.0	201	2 G02312	hypothetical prote
25	7	2.0	201	2 T25706	phenylacrylic acid
26	7	2.0	204	2 G69761	probable oxidoredu
27	7	2.0	204	2 D91120	probable oxidoredu
28	7	2.0	204	2 D85965	hypothetical prote
29	7	2.0	204	2 D81793	hypothetical prote

30	7	2.0	205	2 AH2790	conserved hypothet
31	7	2.0	205	2 G97569	BH278 hypothetica
32	7	2.0	209	2 H65092	hypothetical prote
33	7	2.0	222	2 T31246	hypothetical prote
34	7	2.0	222	2 H75539	branched-chain ami
35	7	2.0	237	2 A75400	probable phosphogl
36	7	2.0	246	2 A84384	hypothetical prote
37	7	2.0	281	2 B83724	transcription regu
38	7	2.0	282	2 I51236	brain-derived neur
39	7	2.0	290	2 B83876	D-alanyl-D-alanine
40	7	2.0	295	2 S35914	regulatory protein
41	7	2.0	295	2 T34774	hypothetical prote
42	7	2.0	298	2 T27209	hypothetical prote
43	7	2.0	300	1 JC5229	protein kinase (EC
44	7	2.0	300	2 AB0762	conserved hypothet
45	7	2.0	300	2 C84853	hypothetical prote

ALIGNMENTS

RESULT 1

T15523

hypothetical protein C16B8.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999

C:Accession: T15523

R:Bentley, D.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C16B8.

A:Reference number: 218365

A:Accession: T15523

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-548 <BEN>

A:Cross-references: EMBL:U41031; NID:G1098982; PID:G1098983; PID:AAA82618.1; CESP:C16

C:Genetic:

A:Gene: CESP:C16B8.1

A:Initons: 13/1; 53/3; 123/1; 176/2; 200/1; 229/3; 255/1; 371/1; 466/3; 512/2

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homo

Query Match Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 AADVMSLGV 249

DB 463 AADVMSLGV 471

RESULT 2

A95405

conserved hypothetical protein SMA2099 [imported] - Sinorhizobium meliloti (strain 102

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: A95405

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bo

proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli

A:Reference number: A95462; PMID:11396509; PMID:11481432

A:Accession: A95405

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-190 <KUR>

A:Cross-references: GB:AB006469; PID:AAK65803.1; PID:G14524305; GSPDB:GN00165

R:Galibert, F.; Finan, T.M.; Long, S.R.; Pulter, A.; Abola, P.; Ampe, F.; Barloy-Hubler

pele, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, B.; Komp, C.; Lelaur

hebaule, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Weller, D.H.; Wong, K.; Yeh,

A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A/Reference number: A96039; MUID:21368234; PMID:11474104
 A/Contents: annotation
 A/Genetics:
 A/Gene: SMA2099
 A/Genome: plasmid

Query Match 2.2%; Score 8; DB 2; Length 190;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 EPARLTA 306
 |||||
 DB 11 EPARLTA 18

RESULT 3

T11854
 protein kinase (EC 2.7.1.-) - *Trypanosoma brucei*

C/Species: *Trypanosoma brucei*
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000

C/Accession: T11854

R/Gale, M.J.; Parsons, M.

Mol. Biochem. Parasitol. 59, 111-122, 1993

A/Title: *Trypanosoma brucei* gene family encoding protein kinases with catalytic domains

A/Reference number: 217363; MUID:93295429; PMID:8515773

A/Accession: T11854

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-431 <GAL>

A/Cross-references: EMBL:L03778; NID:9162169; PID:9162170

A/Genetics:

A/Note: nTKA

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C/Keywords: ATP, phosphotransferase

Query Match 2.2%; Score 8; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ADVMSLCV 249
 |||||
 DB 205 ADVMSLCV 212

RESULT 4

A83266
 conserved hypothetical protein PA3048 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C/Species: *Pseudomonas aeruginosa*
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C/Accession: A83266

R/Slover, C.K.; Pham, X.Q.; Ewain, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,

.i. Loay, S.; Olson, M.V.

Nature 406, 959-964, 2000

A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho

A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: A83266

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-725 <STO>

A/Cross-references: GB:AE004729; GB:AE004091; NID:9949143; PID:AA06436.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Gene: PA3048

Query Match 2.2%; Score 8; DB 2; Length 725;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 339 GLGLDEAR 346
 |||||
 DB 25 GLGLDEAR 32

RESULT 5

AH3258
 cation-transporting ATPase pacc (EC 3.6.1.-) [imported] - *Brucella melitensis* (strain 1

C/Species: *Brucella melitensis*

C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C/Accession: AH3258

R/Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Lo, T.; Ivanova,

.i. Mazur, M.; Goldstein, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lete

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen *Brucella melite*

A/Reference number: AD3252; PMID:11756688

A/Accession: AH3258

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-813 <KUR>

A/Cross-references: GB:AE008917; PID:AAU51235.1; PID:917981923; GSPDB:GN00190

A/Experimental source: strain 16M

C/Genetics:

A/Gene: BMR10053

A/Map position: 1

C/Keywords: hydrolase

Query Match 2.2%; Score 8; DB 2; Length 813;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AAPAGSLT 14
 |||||
 DB 188 AAPAGSLT 195

RESULT 6

T36211
 probable excisionase - *Streptomyces coelicolor* (fragment)

C/Species: *Streptomyces coelicolor*

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000

C/Accession: T36211

R/Oliver, K.; Bentley, D.; Bentley, S.D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, March 1999

A/Reference number: Z21577

A/Accession: T36211

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-33 <OLI>

A/Cross-references: EMBL:AL049573; PIDN:CAB40309.1; GSPDB:GN00070; SCOPDB:SCB39.01c

A/Experimental source: strain A3(2)

C/Genetics:

A/Gene: SCOPDB:SCB39.01c

Query Match 2.0%; Score 7; DB 2; Length 33;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 ATAVATA 65
 |||||
 DB 4 ATAVATA 10

RESULT 7

H44057
 early B3 10.3K protein - human adenovirus 11 (strain BC34) (fragment)

C/Species: *Adenovirus 11* (human adenovirus 11)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 12-Apr-1995

C/Accession: H44057

R/Mel, Y.F.; Wadell, G.

Virology 191, 125-133, 1992

A/Title: The nucleotide sequence of adenovirus type 11 early 3 region: comparison of g

A/Reference number: A44057; MUID:93033102; PMID:1413499

A/Accession: H44057

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-66 <MBI>
A/Cross-references: GB:M94459
C/Superfamily: adenovirus early E3 10.3K protein

Query Match 2.0%; Score 7; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 TAVATAS 66
|||||
DB 22 TAVATAS 28

RESULT 8

T36199
probable excisionase - Streptomyces coelicolor

C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000

R/Saunders, D.C.; Harris, D.; Bentley, S.D.; Parthill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, March 1999
A/Reference number: Z21600

A/Accession: T36199
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-66 <SAL>
A/Cross-references: EMBL:AL035707; PIDN:CAB38895.1; GSPDB:GN00070; SCOPDB:SCB29.20C
A/Experimental source: strain A3(2)
C/Genetics:
A/Genes: SCOPDB:SCB29.20C
C/Superfamily: excisionase

Query Match 2.0%; Score 7; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 ATAVATA 65
|||||
DB 4 ATAVATA 10

RESULT 9

AD3056
hypothetical protein Atu4065 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C/Accession: AD3056
R/Wood, D.W.; Seubald, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, F.; Zhang, S.
Science 294, 2317-2333, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreepan, W.; Perry, M.; Gordon-Kamm,
eter, B.W.

A/Title: The genome of the Natural Genetic Engineering Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193

A/Accession: AD3056
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-72 <KUR>
A/Cross-references: GB:AB08689; PIDN:AL44866.1; PID:G17742513; GSPDB:GN00187
C/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Genes: Atu4065
A/Map position: linear chromosome

Query Match 2.0%; Score 7; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 278 YALPAGL 284
|||||

DB 25 YALPAGL 31

RESULT 10

H89990
hypothetical protein SA1809 [imported] - Staphylococcus aureus (strain N315)

C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C/Accession: H89990
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cu, L.; Og
ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:11418146

A/Accession: H89990
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-80 <KUR>
A/Cross-references: GB:BA000018; PID:G13701796; PIDN:BA843089.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Genes: SA1809

Query Match 2.0%; Score 7; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 KXVLLEN 202
|||||
DB 10 KXVLLEN 16

RESULT 11

S78725
protein YKL053c-a - Yeast (Saccharomyces cerevisiae)

C/Species: Saccharomyces cerevisiae
A/Variety: strain S288C

C/Date: 15-Jan-1999 #sequence_revision 15-Jan-1999 #text_change 19-Apr-2002
C/Accession: S78725; S78074
R/Rasmussen, S.; von Wettstein, D.

submitted to the Protein Sequence Database, March 1994
A/Reference number: S37872
A/Accession: S78725
A/Status: preliminary

A/Molecule type: DNA
A/Residues: 1-86 <RA>
A/Cross-references: EMBL:Z28054; MIPS:YKL053c-a
A/Accession: S78074

A/Molecule type: DNA
A/Residues: 1-86 <RA>
A/Cross-references: EMBL:Z28052; NID:G486071; PID:G1283516; PID:G2980812; MIPS:YKL053c
C/Genetics:
A/Cross-references: SGD:S0007243
A/Map position: 11L
A/Note: YKL053c-a

Query Match 2.0%; Score 7; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 LDERARE 348
|||||
DB 66 LDERARE 72

RESULT 12

BRAD27
early E3 10.3K protein - human adenovirus 3

C/Species: Mastadenovirus h3 (human adenovirus 3)

A/Note: host Homo sapiens (man)
C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 07-Nov-1997
C/Accession: C31830; H29500

R/Signaas, C./ Akusjaer, G., Pettersson, U.

Gene 50, 173-184, 1986

A/Title: Region E3 of human adenoviruses; differences between the oncogenic adenovirus-3

A/Reference number: A91566; MUID:67219876; PMID:3582978

A/Accession: C31830

A/Molecule type: DNA

A/Residues: 1-91 <SIG>

A/Cross-references: GB:M15952; NID:G209901

A/Note: the GenBank entry ADR3AA PID:G209908 differs from the published sequence in hav

C/Superfamily: adenovirus early E3 10.3K protein

C/Keywords: early protein; transmembrane protein

Query Match 2.0%; Score 7; DB 1; Length 91;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 TAVATAS 66

DB 22 TAVATAS 28

RESULT 13

BRAD77

early E3 10.3K protein - human adenovirus 7 (strain Gomen)

C/Species: Mastadenovirus h7 (human adenovirus 7)

C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999

C/Accession: G31830; #52804; C31830; H29500

R/Hong, J.S.; Mullis, K.G.; Engler, J.A.

Virology 167, 545-553, 1988

A/Title: Characterization of the early region 3 and fiber genes of Ad7.

A/Reference number: A94386; MUID:89073758; PMID:2849239

A/Accession: G31830

A/Molecule type: DNA

A/Residues: 1-91 <HON>

A/Cross-references: GB:M23696; NID:G344012; PIDN:AAA53251.1; PID:G576458

R/Kojon, A.B.; Madell, G.

Submitted to the EMBL Data Library, April 1995

A/Description: Sequence analysis of the E3 region and fiber gene of human adenovirus 7h.

A/Reference number: 852798

A/Accession: 852804

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-91 <KAJ>

A/Cross-references: EMBL:Z46954; NID:G762955; PIDN:CAA88813.1; PID:G762962

C/Superfamily: adenovirus early E3 10.3K protein

C/Keywords: early protein; transmembrane protein

Query Match 2.0%; Score 7; DB 1; Length 91;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 TAVATAS 66

DB 22 TAVATAS 28

RESULT 14

QOPSH7

hypothetical protein merr - Pseudomonas aeruginosa transposon Tn501

C/Species: Pseudomonas aeruginosa

C/Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 24-Sep-1999

C/Accession: A04457

R/Misra, T.K.; Brown, N.L.; Pritzinger, D.C.; Pridmore, R.D.; Barnes, W.M.; Haberecht,

Proc. Natl. Acad. Sci. U.S.A. 81, 5975-5979, 1984

A/Title: Mercutio ion-resistance operators of plasmid R100 and transposon Tn501: the begit

A/Reference number: A03556; MUID:85014891; PMID:6091128

A/Accession: A04457

A/Molecule type: DNA

A/Residues: 1-116 <MIS>

A/Cross-references: GB:Z00027; GB:K00031; GB:K01725; GB:K01297; GB:X03406; NID:G43714; F

C/Genetics:

A/Genes: merr

C/Superfamily: merr protein

C/Keywords: transmembrane protein

Query Match 2.0%; Score 7; DB 1; Length 116;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LAVLEPP 111

DB 44 LAVLEPP 50

RESULT 15

B69418

hypothetical protein AF1347 - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C/Accession: B69418

R/Kienk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.B.; Ketchum, K.A.; Dode

, F.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, B.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Moese, C.R.; Venter, J.C.

A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archa

A/Reference number: A69250; MUID:98049343; PMID:9389475

A/Accession: B69418

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-123 <KLE>

A/Cross-references: GB:AB001011; GB:AB000782; NID:G2689334; PIDN:AB89916.1; PID:G2649

Query Match 2.0%; Score 7; DB 2; Length 123;

Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 VLEPPR 78

DB 5 VLEPPR 11

Search completed: January 15, 2004, 15:05:54
Job time : 22 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 15, 2004, 15:05:58 (Search time 37 seconds
(without alignments)
1978.494 Million cell updates/sec)

Title: US-09-909-474d-2

Perfect score: 358

Sequence: 1 MBATPLAPAGSLSRKKRL.....GLGLDEARBEEDREVLVYG 358

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 762491 seqs, 204481190 residues

Word size: 0

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Published Applications_AA.*

1: /cgn2_6/prodata/2/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	274	76.5	358	US-09-799-875-8	Sequence 9, Appl1
2	234	65.4	360	US-10-024-828-9	Sequence 9, Appl1
3	208	58.1	233	US-09-925-301-1102	Sequence 1102, Ap
4	62	17.3	153	US-09-925-301-1367	Sequence 1367, Ap
5	14	3.9	138	US-09-864-761-45767	Sequence 45767, A
6	13	3.6	206	US-10-228-263-2	Sequence 2, Appl1
7	7	1.3	269	US-10-291-172-304	Sequence 304, App
8	9	2.5	97	US-09-864-761-45703	Sequence 45703, A
9	9	2.5	290	US-10-291-172-680	Sequence 680, App
10	9	2.5	548	US-10-369-493-6754	Sequence 6754, Ap
11	8	2.2	159	US-10-156-761-1825	Sequence 12825, A
12	8	2.2	357	US-10-156-761-13581	Sequence 13581, A
13	8	2.2	725	US-09-815-242-5125	Sequence 5125, Ap
14	8	2.2	1352	US-10-156-761-9867	Sequence 9867, Ap
15	8	2.2	1701	US-09-963-959-2	Sequence 2, Appl1

16	7	2.0	8	12	US-10-348-232-168	Sequence 168, App
17	7	2.0	31	12	US-10-029-386-32285	Sequence 32285, A
18	7	2.0	35	9	US-09-864-761-46811	Sequence 46811, A
19	7	2.0	61	12	US-10-029-386-31085	Sequence 31085, A
20	7	2.0	90	15	US-10-083-357-740	Sequence 740, App
21	7	2.0	102	11	US-09-764-891-3529	Sequence 3529, Ap
22	7	2.0	102	11	US-09-764-891-3951	Sequence 3951, Ap
23	7	2.0	114	12	US-10-264-049-3683	Sequence 3683, Ap
24	7	2.0	117	9	US-09-799-875-30	Sequence 30, Appl
25	7	2.0	130	9	US-09-799-875-22	Sequence 22, Appl
26	7	2.0	158	12	US-10-104-047-3610	Sequence 3610, Ap
27	7	2.0	158	12	US-10-108-260A-4566	Sequence 4566, Ap
28	7	2.0	160	11	US-09-984-271-211	Sequence 211, App
29	7	2.0	170	12	US-10-104-047-2785	Sequence 2785, Ap
30	7	2.0	176	9	US-09-799-875-27	Sequence 27, Appl
31	7	2.0	185	15	US-10-156-761-8607	Sequence 8607, Ap
32	7	2.0	186	9	US-09-910-150-14	Sequence 14, Appl
33	7	2.0	201	10	US-09-901-540-3	Sequence 3, Appl1
34	7	2.0	201	10	US-09-896-856-8	Sequence 8, Appl1
35	7	2.0	201	10	US-09-901-257-3	Sequence 3, Appl1
36	7	2.0	201	12	US-10-407-303-8	Sequence 8, Appl1
37	7	2.0	201	15	US-10-107-931-8	Sequence 8, Appl1
38	7	2.0	201	15	US-10-212-793-4	Sequence 4, Appl1
39	7	2.0	212	9	US-09-834-496A-4	Sequence 4, Appl1
40	7	2.0	214	10	US-09-515-806-10	Sequence 10, Appl1
41	7	2.0	214	10	US-09-860-352A-6	Sequence 6, Appl1
42	7	2.0	214	12	US-10-369-493-7805	Sequence 7805, Ap
43	7	2.0	214	12	US-10-410-764-93	Sequence 93, Appl
44	7	2.0	232	11	US-09-809-391-623	Sequence 623, App
45	7	2.0	232	12	US-09-882-171-623	Sequence 623, App

ALIGNMENTS

RESULT 1
US-09-799-875-8
Sequence 8, Application US/09799875
Patient No. US20020034780A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
APPLICANT: Kapeller-Libermann, Rosana
TITLE OF INVENTION: No. US20020034780A1 Human Protein Kinases and Uses
TITLE OF INVENTION: Therefore
FILE REFERENCE: 35800/209996
CURRENT APPLICATION NUMBER: US/09/799, 875
CURRENT FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: 60/182,059
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: 09/659,287
PRIOR FILING DATE: 2000-09-12
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 8
LENGTH: 358
TYPE: PRT
ORGANISM: Homo sapiens
US-09-799-875-8

Query Match 76.5%; Score 274; DB 9; Length 358;
Best Local Similarity 100.0%; Pred. No. 1.5e+247;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

85 ALHCPGTGYTCNVYVQALAVLEBYARLPKHVAPPTVLAGTQLVAFTRTHGM 144
|||||
85 ALHCPGTGYTCNVYVQALAVLEBYARLPKHVAPPTVLAGTQLVAFTRTHGM 144
|||||
145 HSLVSRHRHIPPBEAAVLFROMATLALHCHQGLVLRDILKCPVPADBRKLVLENLE 204
|||||
145 HSLVSRHRHIPPBEAAVLFROMATLALHCHQGLVLRDILKCPVPADBRKLVLENLE 204
|||||
205 DSCVLTGPDSDSLMDKXICPAVYVGPBILSSRASVSGKAADWSLGVALLFTMLAGHYFPQDS 264

[illegible]

```

RESULT 2 --828-9
US-10-024-828-9
/ Sequence 9, Application US/10024828
/ Publication No. US20030036051A1
/ GENERAL INFORMATION:
/ APPLICANT: Virca, Duke
/ APPLICANT: Bird, Timothy A.
/ APPLICANT: Anderson, Dirk M.
/ APPLICANT: Marken, John S.
/ TITLE OF INVENTION: Human CDNA Encoding Polypeptides Having Kinase
/ INVENTOR:

```

CURRENT APPLICATION NUMBER: US/10/024,822-9
 CURRENT FILING DATE: 2001-12-18
 PRIOR APPLICATION NUMBER: US/09/509,902A
 PRIOR FILING DATE: 1999-08-03
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 9
 LENGTH: 360
 TYPE: prt
 ORGANISM: Homo sapiens
 US-10-024-828-9

Query Match 65.4%; Score 234; DB 15; Length 360;
Beet Local Similarity 100.0%; Pred. NC. 3.6e-210;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	85	ALHCPGTGNYCKYVPQVEALAVLEPYARLPNKHVAFTVETAGTOLLVAFPTTHDM	144
Db	127	ALHCPGTGNYCKYVPQVEALAVLEPYARLPNKHVAFTVETAGTOLLVAFPTTHDM	186
QY	145	HSLSVSRRIIEPEBAVLFROMATALACHOHGLVRLDKLCPVADRRKQVLENIIE	204
Db	187	HSLSVSRRIIEPEBAVLFROMATALACHOHGLVRLDKLCPVADRRKQVLENIIE	246
QY	205	DSCTLTGPDDBLMDKHACPAVYGPFIISRSRASYSGKADVMSLGYALTTMLAGHYPPDS	264
Db	247	DSCTLTGPDDBLMDKHACPAVYGPFIISRSRASYSGKADVMSLGYALTTMLAGHYPPDS	306
QY	265	EPVLLFGKIRBGAAVALPAGLSAPARCIVRCLLRRPARRLTATGYLLHPMLROD	318
Db	307	EPVLLFGKIRBGAAVALPAGLSAPARCIVRCLLRRPARRLTATGYLLHPMLROD	360

```

RESULT 3
US-09-925-301-1102
/ Sequence 1102, Application US/09925301
/ Patent No. US20020052308A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA106
/ CURRENT APPLICATION NUMBER: US/09/925,301
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05892
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 1694
/ SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 1102
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1102

```

Query Match	58.1%	Score 208	DB 9	Length 233
Best Local Similarity	100.0%	Pred. NO. 4	8e-166	
Matches 208	0	Mismatches	0	Gaps 0

QY	151	26	211	86	QY	271	146	QY	331	206
RHRIPREAAVLT	PROMTALAHCHQHVLRDLCEPVPADRRKCLVLENTLBSCVLT	RHRIPREAAVLT	PFQMTALAHCHQHVLRDLCEPVPADRRKCLVLENTLBSCVLT	85	QY	211	86	QY	331	206
GPDDSLMDKACAPAYVGP	PIILSRASYSGKADVMSGLVLT	GPDDSLMDKACAPAYVGP	PIILSRASYSGKADVMSGLVLT	145	QY	271	146	QY	331	206
GKTRRGAYVALPAGISAPARCLV	RCLLRBPAPERLTATGILLHPWLRODPMELAPPRSHLM	GKTRRGAYVALPAGISAPARCLV	RCLLRBPAPERLTATGILLHPWLRODPMELAPPRSHLM	205	QY	271	146	QY	331	206
BAAGVVPDGLGLDEAREEGDR	EVLYTG	BAAGVVPDGLGLDEAREEGDR	EVLYTG	358	QY	271	146	QY	331	206
BAAGVVPDGLGLDEAREEGDR	EVLYTG	BAAGVVPDGLGLDEAREEGDR	EVLYTG	233	QY	271	146	QY	331	206

RESULT 4
US-09-925-301-1367
; Sequence 1367, Application US/09925301
; Patent No. US20020052308A1

1 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 2
 3 FILE REFERENCE: PA106
 4
 5 CURRENT APPLICATION NUMBER: US/09/925,301
 6
 7 PRIOR FILING DATE: 2001-08-10
 8
 9 CURRENT APPLICATION NUMBER: PCT/US00/05882
 10
 11 PRIOR FILING DATE: 2000-03-08
 12
 13 PRIOR APPLICATION NUMBER: 60/124,270
 14
 15 PRIOR FILING DATE: 1999-03-12
 16
 17 NUMBER OF SEQ ID NOS: 1694
 18
 19 SOFTWARE: Patent Ver. 2.0

?	TYPE: PRT	
?	ORGANISM: Homo sapiens	
?	FEATURE:	
?	NAME/KEY: SITE	
?	LOCATION: (136)	
?	OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids	
?	NAME/KEY: SITE	
?	LOCATION: (138)	
?	OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids	
?	NAME/KEY: SITE	
?	LOCATION: (141)	
?	OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids	
?	NAME/KEY: SITE	
?	LOCATION: (142)	
?	OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids	
?	NAME/KEY: SITE	
?	LOCATION: (143)	
?	OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids	
?	NAME/KEY: SITE	
?	LOCATION: (152)	
?	OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids	
?	IS-09-925-301-1367	

Query Match	17.3%	Score 62	DB 9	Length 153
Best Local Similarity	100.0%	Pred. No.	9.3e-50	
Matches	62	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0
Q1	1	MRATPLAAPXGJSRKKRLDNDLDRTEVQKRAKSGPQPRPLPCLLLPSPPTAPDRAT	60	

Db 74 MRATPLAPAGSLSKKRLBDNDLTERPVQKARSQGPLPCLPLSPPTAPDRAT 133
Qy 61 AV 62
134 AV 135

RESULT 5
US-09-864-761-45767
Sequence 45767, Application US/09864761
Patent No. US20020046763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmlca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 45767
LENGTH: 138
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009486.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
OTHER INFORMATION: SWISSPROT HIT: 074536, EVALUATE 1.00e-18

OTHER INFORMATION: EST_HUMAN HIT: BE897149.1, EVALUATE 2.00e-71
US-09-864-761-45767

Query Match 3.9%; Score 14; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 6,1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 236 YSGKADVMSLGV 249
Db 41 YSGKADVMSLGV 54

RESULT 6
US-10-228-263-2
Sequence 2, Application US/10228263
Publication No. US2003009985A1
GENERAL INFORMATION:
APPLICANT: Li, Jing
TITLE OF INVENTION: AMPLIFIED GENE INVOLVED IN CANCER
FILE REFERENCE: 38002-0034
CURRENT FILING DATE: 2002-12-04
PRIOR APPLICATION NUMBER: US 60/330,797
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: US 60/314,555
PRIOR FILING DATE: 2001-08-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 206
TYPE: PRT
ORGANISM: Homo sapiens
US-10-228-263-2

Query Match 3.6%; Score 13; DB 15; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.00075;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 237 YSGKADVMSLGV 249
Db 94 YSGKADVMSLGV 106

RESULT 7
US-10-291-172-304
Sequence 304, Application US/10291172
Publication No. US20030228584A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: No. US20030228584A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-045
CURRENT FILING DATE: US/10/291,172
CURRENT FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/693,267
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/665,363
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/616,847
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 09/596,193
PRIOR FILING DATE: 2000-06-17
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 752
SEQ ID NO 304
LENGTH: 269
TYPE: PRT
ORGANISM: Homo sapiens
US-10-291-172-304
Query Match 3.6%; Score 13; DB 12; Length 269;

Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 237 YSGKADVMSIGV 249
DB 157 YSGKADVMSIGV 169

RESULT 8
US-09-864-761-45703
Sequence 45703, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR
FILB REFERENCE: Aecm1ca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Lasting Engine vers. 1.1
SEQ ID NO 45703
LENGTH: 97
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009486.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.62
OTHER INFORMATION: SWISSPROT HIT: P57058, EVALU 1.00e-06

OTHER INFORMATION: EST_HUMAN HIT: AUI27403.1, EVALU 4.00e-51
US-09-864-761-45703
Query Match 2.5%; Score 9; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 GLVLRDLKL 185
DB 80 GLVLRDLKL 88

RESULT 9
US-10-291-172-680
Sequence 680, Application US/10291172
Publication No. US20030228584A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NO. US20030228584A1 Nucleic Acids and Polypeptides
FILB REFERENCE: 21272-045
CURRENT FILING DATE: US/10/291,172
CURRENT FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/693,267
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/665,363
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/616,847
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 09/596,193
PRIOR FILING DATE: 2000-06-17
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 752
SEQ ID NO 680
LENGTH: 290
TYPE: PRT
ORGANISM: Homo sapiens
US-10-291-172-680

Query Match 2.5%; Score 9; DB 12; Length 290;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 237 YSGKADVMSIGV 245
DB 168 YSGKADVMSIGV 176

RESULT 10
US-10-369-493-6754
Sequence 6754, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILB REFERENCE: 38-10(52052)B
CURRENT FILING DATE: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 6754
LENGTH: 548
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-6754

Query Match 2.5% Score 9: DB 12: Length 548;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 AADVWSLGV 249
|||||
Db 463 AADVWSLGV 471

RESULT 11
US-10-156-761-12825
; Sequence 12825, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12825
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12825

Query Match 2.2% Score 8: DB 15: Length 159;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 338 DGLGLDBA 345
|||||
Db 125 DGLGLDBA 132

RESULT 12
US-10-156-761-13581
; Sequence 13581, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13581
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13581

Query Match 2.2% Score 8: DB 15: Length 357;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 246 SIGVALFT 253
|||||
Db 108 SIGVALFT 115

RESULT 13
US-09-815-242-5125
; Sequence 5125, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5125
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5125

Query Match 2.2% Score 8: DB 9: Length 725;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 339 GGLGDBAR 346
|||||
Db 25 GGLGDBAR 32

RESULT 14
US-10-156-761-9867
; Sequence 9867, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761

```

/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 9867
/ LENGTH: 1352
/ TYPE: PRT
/ ORGANISM: Streptomyces avermitilis
US-10-156-761-9867

```

```

Query Match      2.2%; Score 8; DB 15; Length 1352;
Best Local Similarity 100.0%; Pred.No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      301 ABLTATG 308
      |||||
Db      846 ABLTATG 853

```

```

RESULT 15
US-09-963-959-2
/ Sequence 2, Application US/09963959
/ Patent No. US20020165145A1
/ GENERAL INFORMATION:
/ APPLICANT: Meyers, Rachel A.
/ TITLE OF INVENTION: 33521, A NOVEL HUMAN GUANINE NUCLEOTIDE EXCHANGE FAMILY MEMBER
/ TITLE OF INVENTION: AND USES THEREOF
/ FILE REFERENCE: 10448-095001
/ CURRENT APPLICATION NUMBER: US/09/963,959
/ CURRENT FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: 60/235,033
/ PRIOR FILING DATE: 2000-09-25
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: PaetSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 1701
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-963-959-2

```

```

Query Match      2.2%; Score 8; DB 10; Length 1701;
Best Local Similarity 100.0%; Pred.No. 2.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      11 GLSRRKR 18
      |||||
Db      388 GLSRRKR 395

```

```

Search completed: January 15, 2004, 15:11:57
Job time : 38 secs

```


XX Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections -
 XX
 PS Claim 7, Figure 2, 433pp, English.
 XX
 CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
 CC novel protein kinases have been identified as members of the tyrosine
 CC or serine/threonine kinase (PTK and STR) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancer (especially cancers of hematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.
 XX
 SQ Sequence 358 AA;

Query Match 76.5%; Score 274; DB 22; Length 358;
 Best Local Similarity 100.0%; Pred. No. 4e-258;
 Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 ALHCPGTGYTCYRYVQBALAVLEPYARLPKHVARPTEVLAQTOLLYAFPTRTGDM 144
 DB 85 ALHCPGTGYTCYRYVQBALAVLEPYARLPKHVARPTEVLAQTOLLYAFPTRTGDM 144
 QY 145 HSLVRSRHRIRPEBAANTLPRQMTALAHCHQHGVLVDLKLCPVFADEBRKKLVLENLE 204
 DB 145 HSLVRSRHRIRPEBAANTLPRQMTALAHCHQHGVLVDLKLCPVFADEBRKKLVLENLE 204
 QY 205 DSCVLGTPDDSLMDKACPAVYGPBEILSSRASYSQKADVWSLGVALLFTMLAGHYPPQDS 264
 DB 205 DSCVLGTPDDSLMDKACPAVYGPBEILSSRASYSQKADVWSLGVALLFTMLAGHYPPQDS 264
 QY 265 EPTVLFGKIRRGAYALPAGLSAPARCLVRLRREPERLTATGILLHFWLRDPMPLAP 324
 DB 265 EPTVLFGKIRRGAYALPAGLSAPARCLVRLRREPERLTATGILLHFWLRDPMPLAP 324
 QY 325 TRSHLWBAQVVPDGLDEARBBEGDREVLYG 358
 DB 325 TRSHLWBAQVVPDGLDEARBBEGDREVLYG 358

RESULT 2
 AAM38908
 ID AAM38908 standard; Protein; 358 AA.

XX AAM38908.

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2053.

XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
 XX leukemia.

XX Homo sapiens.

XX MO20015312-A1.

PD 26-UTL-2001.
 PF 26-DEC-2000; 2000MO-US34263.
 XX
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-UTL-2000; 2000US-0598042.
 PR 19-UTL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren P, Wang D;
 PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI, 2001-442253/47.
 DR N-PSDB; AA158064.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 3; SEQ ID NO 2053; 10078bp, English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nerve injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

QY 98 VYVQBALAVLEPYARLPKHVARPTEVLAQTOLLYAFPTRTGDMHSLVRSRHRIRPEP 157
 DB 98 VYVQBALAVLEPYARLPKHVARPTEVLAQTOLLYAFPTRTGDMHSLVRSRHRIRPEP 157
 QY 158 EAAVLFQMAATLALAHCHQHGVLVDLKLCPVFADEBRKKLVLENLEDSCLVGPDDSLW 217
 DB 158 EAAVLFQMAATLALAHCHQHGVLVDLKLCPVFADEBRKKLVLENLEDSCLVGPDDSLW 217
 QY 218 DKACAPAYVGPBEILSSRASYSQKADVWSLGVALLFTMLAGHYPPQDSBPVLFGKIRGA 277
 DB 218 DKACAPAYVGPBEILSSRASYSQKADVWSLGVALLFTMLAGHYPPQDSBPVLFGKIRGA 277
 QY 278 YALPAGLSAPARCLVRLRREPERLTATGILLHFWLRDPMPLAPTRSHLWBAQVVP 337
 DB 278 YALPAGLSAPARCLVRLRREPERLTATGILLHFWLRDPMPLAPTRSHLWBAQVVP 337
 QY 338 DGLGLDEARBBEGDREVLYG 358
 DB 338 DGLGLDEARBBEGDREVLYG 358

RESULT 3
 AAM40694

ID	AA040694 standard; Protein; 393 AA.
AC	AA040694;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 5625.
XX	
KW	Human; nocrotic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nerve system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Dreger Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
KM	
XX	
OS	Homo sapiens.
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000MO-US34263.
PR	21-JAN-2000; 2000US-0488725.
XX	
PR	25-APR-2000; 2000US-0552317.
XX	
PR	09-JUL-2000; 2000US-0598042.
XX	
PR	19-JUL-2000; 2000US-0620312.
XX	
PR	03-AUG-2000; 2000US-0653450.
XX	
PR	14-SEP-2000; 2000US-0662191.
XX	
PR	19-OCT-2000; 2000US-0693036.
XX	
XX	29-NOV-2000; 2000US-0727344.
PA	(HSE-) HSEQ INC.
XX	
P1	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
P1	Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
P1	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
DR	N-PBDB; AAI59850.
XX	
XX	MP1: 2001-44253/47.
PT	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -
PS	Example 2; SEQ ID NO 5625; 10078pp; English.
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AA036642-AA042213) with nocrotic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Dreger Syndrome. Other uses include the utilization of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukemias and C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed specification.
CC	
CC	
XX	
Seq	Sequence 393 AA:
Query Match	72.9%; Score 261; DB 22; Length 393;
Best Local Similarity	100.0%; Pred. No. 2e-245;
Matches 261; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
98	VYPOEALAVIEPVARLPKHGVAPPTVLGTOILVAFPTTHDDMSLVSRRIIEP 157
133	VYPOEALAVIEPVARLPKHGVAPPTVLGTOILVAFPTTHDDMSLVSRRIIEP 192

QY	156	EAAVLFROMATLACHOHGVLRLDKCRFPAPRRKKVLLENLSDSCVLTGPDDSLW	217
Db	193	EAAVLFROMATLACHOHGVLRLDKCRFPADRRKKVLLENLSDSCVLTGPDDSLW	252
QY	218	DKHACPAYVGPETLSSRASYSYGKADWWSLGVALFTMLAGHYPFQDSBPVLLFGKIRGA	277
Db	253	DGHACPAYVGPETLSSRASYSYGKADWWSLGVALFTMLAGHYPFQDSBPVLLFGKIRGA	312
QY	278	YALPAGLSAPARCLVRCILRRRPAERLTATGILLHPWLRODPMPLAFTSHLWBAQQVP	337
Db	313	YALPAGLSAPARCLVRCILRRRPAERLTATGILLHPWLRODPMPLAFTSHLWBAQQVP	372
QY	338	DGLGIDEAREBERGDEEVLYG	358
Db	373	DGLGIDEAREBERGDEEVLYG	393
RESULT 4			
AA69157	AA69157 standard; Peptide; 360 AA.		
XX	AA69157/		
XX	30-MAY-2000	(first entry)	
XX	PeptideA JJ503-KS comprising domain VIA to XI of a protein kinase.		
XX	Kinase activity; molecular weight marker; isoelectric focusing marker;		
XX	peptide fragmentation control; cellular signal transduction.		
XX	Homo sapiens.		
XX	MO200008180-A2.		
XX	17-FEB-2000.		
XX	03-AUG-1999;	99WO-US17630.	
XX	04-AUG-1998;	98US-0095270.	
XX	11-SEP-1998;	98US-0099972.	
XX	(IMMV) IMMUNEX CORP.		
XX	Vitrea GD, Bird TA, Anderson DM, Marken JS;		
XX	WPI; 2000-195584/17.		
XX	N-PSDB; AAZ61155.		
XX	New human kinase polypeptides and polynucleotides used as molecular		
XX	weight markers and as controls for peptide fragmentation		
XX	Claim 2; Page 10; 60pp; English.		
XX	The present sequence represents a partial polypeptide which has kinase		
XX	activity. The kinase polynucleotides can be used to express the		
XX	polypeptides, and as probes to identify nucleic acids encoding		
XX	proteins having kinase activity. The kinase polypeptides and		
XX	fragmented polypeptides are used as molecular weight and isoelectric		
XX	focusing markers, and as controls for peptide fragmentation. They also		
XX	have a number of therapeutic uses as kinases play a central role in		
XX	cellular signal transduction. The polypeptides could also be used to		
XX	identify binding partner proteins. The polypeptides can also be used as		
XX	a reagent to identify any proteins that the polypeptide regulates, and		
XX	as proteins with which it might interact. The polypeptides may also be		
XX	used for preparation of antibodies. The antibodies can be used in		
XX	assays to detect the presence of the protein, and to purify the protein		
XX	by immunoinfinity chromatography.		
XX	Sequence	360 AA;	
XX	Query Match	65.4%;	Score 234; DB 21; Length 360;
XX	Best Local Similarity	100.0%;	Pred. No. 3.6e-219;
XX	Matches 234;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;

QY 85 ALHCPGTGYTCYKYPVQEBALVLEPYARLPKHVAPTEVLAQTOLLYAFPTTTHGDM 144
 DB 127 ALHCPGTGYTCYKYPVQEBALVLEPYARLPKHVAPTEVLAQTOLLYAFPTTTHGDM 186
 QY 145 HSLVSRRIPEPEBAVLPROMATLALHCHQGLVLRDLKLCRFVPADRERKKVLBNLE 204
 DB 187 HSLVSRRIPEPEBAVLPROMATLALHCHQGLVLRDLKLCRFVPADRERKKVLBNLE 246
 QY 205 DSCVLTGPDSDLMDKHAFCPAYVGPBILSSRASYSCKADVWSLGVALLFTMLAGHYPPDS 264
 DB 247 DSCVLTGPDSDLMDKHAFCPAYVGPBILSSRASYSCKADVWSLGVALLFTMLAGHYPPDS 306
 QY 265 EPTVLTGKIRGAYALPAGLSAPARCLVRCILRRPAPRLTATGILLHPMLRQD 318
 DB 307 EPTVLTGKIRGAYALPAGLSAPARCLVRCILRRPAPRLTATGILLHPMLRQD 360

RESULT 5
 AAB43657
 ID AAB43657 standard; Protein, 233 AA.
 AC AAB43657;
 XX
 DT 08-FEB-2001 (first entry)
 DE Human cancer associated protein sequence SEQ ID NO:1102.
 XX
 KM Human; cancer associated gene; cancer antigen; detection; cancer;
 KM diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
 KM antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
 KM antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
 KM dermatologic; neuroprotective; thrombolytic; coagulant; nocotropic;
 KM vasotropic; antipsoriatic; antidiabetic; gene therapy; inflammation;
 KM immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KM allergic reaction; graft versus host disease; organ rejection;
 KM haemostatic; thrombolytic; cardiovascular disorder; infection;
 KM neurological disease; drug screening.
 KM
 OS Homo sapiens.
 XX
 PN MO20055350-A1.
 PD 21-SEP-2000.
 XX
 PP 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM,
 XX
 DR WPI, 2000-587533/55.
 DR N-P8DB; AAC77866.
 XX
 FT Novel isolated nucleic acids comprising sequences encoding peptides
 FT useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 11, Page 1707-1708; 2352pp; English.
 XX
 CC AAC7667 to AAC76448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerability; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
 CC antiinflammatory; antihypertoid; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nocotropic; vasotropic; antipsoriatic and antidiabetic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating

CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptide, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC76449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 233 AA;
 Query Match 58.1%; Score 208; DB 21; Length 233;
 Beest Local Similarity 100.0%; Pred. No. 5.2e-194; Indels 0; Gaps 0;
 Matches 208; Conservative 0; Mismatches 0

QY 151 RHRIPPEBAVLPROMATLALHCHQGLVLRDLKLCRFVPADRERKKVLBNLEDSCVLT 210
 DB 26 RHRIPPEBAVLPROMATLALHCHQGLVLRDLKLCRFVPADRERKKVLBNLEDSCVLT 85
 QY 211 GPDSLMDKHAFCPAYVGPBILSSRASYSCKADVWSLGVALLFTMLAGHYPPDSBPVLLF 270
 DB 86 GPDSLMDKHAFCPAYVGPBILSSRASYSCKADVWSLGVALLFTMLAGHYPPDSBPVLLF 145
 QY 271 GKIRGAYALPAGLSAPARCLVRCILRRPAPRLTATGILLHPMLRQDMPPLPTRSHLM 330
 DB 146 GKIRGAYALPAGLSAPARCLVRCILRRPAPRLTATGILLHPMLRQDMPPLPTRSHLM 205
 QY 331 EAAQVPPDGLGLDEAREEGDREVVLYG 358
 DB 206 EAAQVPPDGLGLDEAREEGDREVVLYG 233

RESULT 6
 AAB20326
 ID AAB20326 standard; Protein, 358 AA.
 XX
 AC AAB20326;
 XX
 DT 29-MAY-2001 (first entry)
 DE Human protein phosphatase and kinase protein-5.
 XX
 DS Human protein phosphatase and kinase protein-5.
 XX
 KM Protein phosphatase and kinase protein; PPHK-5; human;
 KM gastrointestinal disorder; immune system disorder;
 KM neurological disorder; cell proliferative disorder; cancer;
 KM diagnosis; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Region
 FT 142..315
 FT /note= "protein kinase signature"
 FT Domain
 FT 242..264
 FT /note= "tyrosine kinase catalytic domain signature"
 FT Domain
 FT 78..303
 FT /note= "protein kinase domain signature"
 FT Modified-site
 FT 54
 FT /note= "O-phosphorylated"
 FT Modified-site
 FT 90
 FT /note= "O-phosphorylated"
 FT Modified-site
 FT 140
 FT /note= "O-phosphorylated"
 FT Modified-site
 FT 210
 FT /note= "O-phosphorylated"
 FT Modified-site
 FT 215
 FT /note= "O-phosphorylated"
 FT Modified-site
 FT 14
 FT /note= "O-phosphorylated"
 FT Modified-site
 FT 27
 FT /note= "O-phosphorylated"
 FT Modified-site
 FT 95
 FT /note= "O-phosphorylated"
 FT Modified-site
 FT /note= "O-phosphorylated"

```

FT Modified-site 232
FT /note="O-phosphorylated"
FT Modified-site 238
FT /note="O-phosphorylated"
XX
XX WO200120004-A2.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000MO-US25515.
XX
XX 15-SEP-1999; 99US-0154141.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,
XX Lu DM;
XX
XX WPI; 2001-244811/25.
XX
XX N-PSDB; AAF30480.
XX
XX
XX Novel human protein phosphatase and kinase proteins for diagnosis,
XX treatment and prevention of gastrointestinal, immune system,
XX neurological and cell proliferative disorders -
XX
XX Claim 1; Page 88-89; 103pp; English.
XX
XX The present sequence is that of novel human protein phosphatase
XX and kinase protein PPHK-5, as predicted from Incyte Clone ID No.
XX 127150CB1 (see AAF30480). Tissues that express PPHK-5 (as a
XX fraction of total tissues expressing PPHK-5) include reproductive
XX (0.288), gastrointestinal (0.212) and haematopoietic or immune
XX (0.192). Diseases or conditions associated with tissues expressing
XX PPHK-5 (as a fraction of total tissues expressing PPHK-5) include
XX cancer (0.577), inflammation or trauma (0.327) and cell
XX proliferation (0.308). The encoded protein shows homology to rat
XX kinase. The invention provides human PPHK-1 to -11 polypeptides
XX (see AAB20322-32) and polynucleotides (see AAF30476-86). It also
XX provides expression vectors, host cells, antibodies, agonists and
XX antagonists, as well as methods for diagnosing, treating or
XX preventing disorders associated with expression of PPHK, including
XX gastrointestinal disorders, immune system disorders, neurological
XX disorders and cell proliferative disorders, including cancer.
XX
XX Sequence 358 AA;
XX
XX Query Match 58.1%; Score 208; DB 22; Length 358;
XX Best Local Similarity 100.0%; Pred. No. 7.3e-154;
XX Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 151 RHRIPEPAAYLFRQMATALAHCHQHGLVLDLKLRCFVPADREKLVLENTLEDSCVLT 210
XX |||||
XX DB 151 RHRIPEPAAYLFRQMATALAHCHQHGLVLDLKLRCFVPADREKLVLENTLEDSCVLT 210
XX
XX QY 211 GPDDSLMDKACPAVVGPEILSSRASYSGKADVMSLGVALLFTMLAGHYPPQDSBPVLLF 270
XX |||||
XX DB 211 GPDDSLMDKACPAVVGPEILSSRASYSGKADVMSLGVALLFTMLAGHYPPQDSBPVLLF 270
XX
XX QY 271 GKIRBGAYALPAGISAPARCLVRCLLRREPARELTATGIIILHPMLRQDPMPLAPRRSLTM 330
XX |||||
XX DB 271 GKIRBGAYALPAGISAPARCLVRCLLRREPARELTATGIIILHPMLRQDPMPLAPRRSLTM 330
XX
XX QY 331 EAAQVVPDGLGLDEAREEGDREVVLYG 358
XX |||||
XX DB 331 EAAQVVPDGLGLDEAREEGDREVVLYG 358
XX
XX RESULT 7
XX ABB80976
XX ID ABB80976 standard; Protein; 360 AA.
XX AC
XX ABB80976;

```

```

DT 21-OCT-2002 (first entry)
XX
XX Human tribbles homologue-3 (htrb-3) polypeptide encoding DNA.
XX
XX KW Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; Ap-1;
XX KW antirheumatic; antidiabetic; antipneumatic; osteoparatic;
XX KW osteoporotic; cardiac; cytosolic; haemostatic; immunosuppressive;
XX KW antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
XX KW tumour necrosis factor; TNF; htrb-3.
XX
XX OS Homo sapiens.
XX
XX PN WO200253743-A2.
XX
XX PD 11-JUL-2002.
XX
XX PF 08-JAN-2002; 2002MO-US00070.
XX
XX PR 08-JAN-2001; 2001US-260294P.
XX
XX PA (INTE-) INTERLEUKIN GENETICS INC.
XX
XX PI Dower S, Quantstrom E, Kiss-Toth E;
XX
XX DR WPI; 2002-590635/63.
XX
XX N-PSDB; AAB86479.
XX
XX
XX Novel isolated human tribbles homologue-1 polypeptide for inhibiting
XX Ap-1-mediated inflammatory signal in a cell, and activating
XX ERK-mediated signal e.g. Ap-1-mediated gene activation signal in a cell
XX
XX Example 10; Fig 11B; 131pp; English.
XX
XX The invention provides an isolated human tribbles homologue-1 (htrb-1,
XX also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
XX htrb polypeptide is useful for inhibiting an Ap-1 mediated inflammatory
XX signal in a cell. The polypeptide employed in the method is preferably
XX htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or
XX htrb-3 N C. It is also useful for providing htrb agonist activity for
XX activating an ERK-mediated signal e.g. Ap-1-mediated gene activation
XX signal, an estrogen receptor-mediated gene activation signal, an
XX fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
XX in a cell. Herb modulators are useful for modulating Ap-1 mediated
XX inflammatory signal in a cell such as tumor necrosis factor (TNF)
XX induced inflammatory signal, or an interleukin induced inflammatory
XX signal. Herb proteins are useful in screening assays, predictive medicine
XX and in therapeutics or prophylactics. The htrb proteins are useful for
XX screening compounds e.g. for treating and/or preventing diseases caused
XX by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
XX psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
XX and cancers. The herb therapeutics are useful for antagonizing
XX interleukin-1 dependent disorders of human placenta, intraventricular
XX hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
XX and inflammation or autoimmune disorders. The present sequence represents
XX the htrb-3 polypeptide.
XX
XX Sequence 360 AA;
XX
XX Query Match 45.5%; Score 163; DB 23; Length 360;
XX Best Local Similarity 100.0%; Pred. No. 4.8e-150;
XX Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 196 KCVLENTLEDSCVLTGPDDSLMDKACPAVVGPEILSSRASYSGKADVMSLGVALLFTML 255
XX |||||
XX DB 196 KCVLENTLEDSCVLTGPDDSLMDKACPAVVGPEILSSRASYSGKADVMSLGVALLFTML 257
XX
XX QY 256 AGHYPPQDSBPVLLFGKIRBGAYALPAGISAPARCLVRCLLRREPARELTATGIIILHPML 315
XX |||||
XX DB 256 AGHYPPQDSBPVLLFGKIRBGAYALPAGISAPARCLVRCLLRREPARELTATGIIILHPML 317
XX
XX QY 316 RODPMPLAPRRSLTMEAAQVVPDGLGLDEAREEGDREVVLYG 358
XX |||||

```

DB 318 RODMPLAFTRSHLWEAAQVVPDGLGLDEARREBEGDREVLYG 360

RESULT 8
AAB85791
ID AAB85791 standard; Protein; 323 AA.

AC AAB85791;
XX
XX
XX 29-OCT-2001 (first entry)
XX
XX Human kinase PKIN-10.
XX
XX PKIN, kinase, cytosolic; immunosuppressive; immunostimulant; human;
XX antiarteriosclerotic; cardiac; gene therapy; antisense therapy.
XX
XX Homo sapiens.
XX
XX WO200160991-A2.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05240.
XX
XX 17-FEB-2000; 2000US-0183582.
XX 02-MAR-2000; 2000US-0186559.
XX 09-MAR-2000; 2000US-0188606.
XX 17-MAR-2000; 2000US-0189998.
XX 30-MAR-2000; 2000US-0193851.
XX
XX (INCYTE GENOMICS INC.
XX
XX Tang YT, Buford N, Gandhi AR, Patterson C, Khan PA, Yue H,
XX PI Hafalia A, Shih LL, Tribouley CM, Yao MG, Burrill JD, Marcus GA,
XX PI Zingler KA, Lu DM, Bandman O, Policky JL, Griffin JA, Thornton M,
XX PI Nguyen DB, Lai P, Walsh RT,
XX
XX WPI; 2001-514771/56.
XX DR N-PSDB; AAB76218.
XX
XX Isolated human kinase polypeptides useful in the diagnosis, treatment
XX PT and prevention of cancer, immune disorders and disorders affecting
XX PT growth and development -
XX
XX Claim 1; Page 115; 126pp; English.

CC The invention provides human kinases (PKIN) and polynucleotides encoding
CC PKIN. The PKIN polypeptides can be expressed using standard recombinant
CC methodology. The PKIN polypeptides, polynucleotides, modulators and
CC specific antibodies are useful in the diagnosis, treatment and prevention
CC of cancer, immune disorders, disorders affecting growth and development,
CC atherosclerosis, and other cardiovascular diseases, and lipid disorders
CC and in the assessment of the effects of exogenous compounds on the
CC expression of nucleic acid sequences of human kinases. The present
CC sequence represents a human PKIN-10 polypeptide.
XX
XX Sequence 323 AA;
SQ

Query Match 32.1%; Score 115; DB 22; Length 323;
Best Local Similarity 100.0%; Pred. No. 2,4e-103;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 VNSLGVALTFTMLAGHPFQDSSEVLLFGKIRGAYVLLPGLSPARCLVRCILRRPAR 303
DB 209 VNSLGVALTFTMLAGHPFQDSSEVLLFGKIRGAYVLLPGLSPARCLVRCILRRPAR 268
QY 304 LTAATGILHPMLRQDMPPLAFTRSHLWEAAQVVPDGLGLDEARREBEGDREVLYG 358
DB 269 LTAATGILHPMLRQDMPPLAFTRSHLWEAAQVVPDGLGLDEARREBEGDREVLYG 323

RESULT 9
AAB06093

ID AAB06093 standard; Protein; 278 AA.

AC AAB06093;
XX
XX
XX 10-MAY-2002 (first entry)
XX
XX Human NS protein sequence SEQ ID NO:185.
XX
XX
XX Human, cytosolic; osteopathic; gynaecological; neuroprotective;
XX antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
XX vasoregic; antiarteriosclerotic; antiinflammatory; dermatological;
XX anorectic; muscular; antiferility; cardiovascular; anticoagulant;
XX antifibrinolytic; hypotension; antiseismic; immunomodulator; cardiac;
XX gastrointestinal; vitruclide; antilucer; cerebroprotective; noctropic;
XX contraceptive; vaccine; gene therapy; cancer; osteoporosis; dysomnia;
XX endometriosis; degenerative disease; multiple sclerosis; psoriasis;
XX rheumatoid arthritis; cataract; reestenosis; atherosclerosis; glaucoma;
XX inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
XX infertility; cardiovascular disease; coagulation disease; hypertension;
XX ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
XX diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
XX gastric ulcer; Alzheimer's disease.
XX
XX Homo sapiens.
XX
XX WO200206315-A2.
XX
XX 24-JAN-2002.
XX
XX 17-JUL-2001; 2001WO-IL00653.
XX
XX 18-JUL-2000; 2000IL-0137345.
XX 15-DEC-2000; 2000IL-0140354.
XX
XX (COMP-) COMPUGEN LTD.
XX
XX Mintz L, Freilich S, Bernstein J;
XX
XX WPI; 2002-155037/20.
XX DR N-PSDB; ABL39747.
XX
XX One hundred and twenty eight novel nucleic acid sequences, useful for
XX PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
XX Claim 6; Page 213-214; 290pp; English.

CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06093 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, vitruclide,
CC vasoregic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiferility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiseismic, cardiac,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilizer, antilucer,
CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
CC noctropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dysomnia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, reestenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
XX Sequence 278 AA;
SQ

Query Match 23.2%; Score 83; DB 23; Length 278;
Best Local Similarity 100.0%; Pred. No. 3.1e-72;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSKKRLTLDNDTERPVQKARSGPOPRLPCLPLSPPTAPDRAT 60
 DB 29 MRATPLAAPAGSLSKKRLTLDNDTERPVQKARSGPOPRLPCLPLSPPTAPDRAT 88
 QY 61 AVATASRLGPPVYLLEPEEGRAY 83
 DB 89 AVATASRLGPPVYLLEPEEGRAY 111

RESULT 10
 ID ABB06108 standard; Protein; 278 AA.

XX ABB06108;
 DT 10-MAY-2002 (first entry)
 XX Human NS protein sequence SEQ ID NO:200.

XX Human; cytoostatic; osteopathic; gynaecological; neuroprotective;
 KM antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KM vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
 KM anorectic; muscular; antinfertility; cardiovascular; anticoagulant;
 KM antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiant;
 KM anticonvulsant; antidiabetic; tranquilliser; antidepressant; neurolitic;
 KM gastrointestinal; virucide; antitumor; cerebroprotective; nootropic;
 KM contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KM endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KM rheumatoid arthritis; cataract; resection; atherosclerosis; glaucoma;
 KM inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KM infertility; cardiovascular disease; coagulation disease; hypertension;
 KM ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KM diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KM gastric ulcer; Alzheimer's disease.

XX Homo sapiens.
 OS
 XX MO300206315-A2.
 XX 24-JAN-2002.
 XX 17-JUL-2001; 2001WO-1L00653.
 XX 18-JUL-2000; 2000IL-0137345.
 PR 15-DEC-2000; 2000IL-0140354.
 XX (COMP-) COMPUGEN LTD.
 XX Mintz L, Freilich S, Bernstein J;
 PI WPI, 2002-155037/20.
 DR N-PSDB; ABL39762.
 XX One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
 XX Claim 6; Page 231-232; 290pp; English.

XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06144. The novel sequences
 CC (NS) can have cytoostatic, osteopathic, gynaecological, neuroprotective,
 CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
 CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, anti-HIV, antinfertility, cardiovascular,
 CC anticonvulsant, antifibrinolytic, hypotension, antiaesthetic, cardiant,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antitumor,
 CC antidiabetic, gastrointestinal, aeurolitic, cerebroprotective,
 CC nootropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antitense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,

CC cataracts, resection, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive.

XX
 SQ Sequence 278 AA;
 Query Match 23.2%; Score 83; DB 23; Length 278;
 Best Local Similarity 100.0%; Pred. No. 3.1e-72;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSKKRLTLDNDTERPVQKARSGPOPRLPCLPLSPPTAPDRAT 60
 DB 29 MRATPLAAPAGSLSKKRLTLDNDTERPVQKARSGPOPRLPCLPLSPPTAPDRAT 88
 QY 61 AVATASRLGPPVYLLEPEEGRAY 83
 DB 89 AVATASRLGPPVYLLEPEEGRAY 111

RESULT 11
 ID AAB43922 standard; Protein; 153 AA.

XX AAB43922;
 DT 08-FEB-2001 (first entry)
 XX Human cancer associated protein sequence SEQ ID NO:1367.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KM diagnosis; cytoostatic; proliferative; vunerary; immunomodulator;
 KM antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
 KM antiinflammatory; antihypertensive; antiallergic; antibacterial; cardiant;
 KM dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KM vasotropic; antipsoriatic; antitumor; gene therapy; inflammation;
 KM immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KM allergic reaction; graft versus host disease; organ rejection;
 KM haemostatic; thrombolytic; cardiovascular disorder; infection;
 KM neurological disease; drug screening.

XX Homo sapiens.
 OS
 XX WO20005350-A1.
 XX 21-SEP-2000.
 XX 08-MAR-2000; 2000WO-US05882.
 XX 12-MAR-1999; 99US-0124270.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM;
 PI WPI, 2000-587533/55.
 DR N-PSDB; AAC78131.
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 XX Claim 11; Page 2031; 2352pp; English.

XX AAC7607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytoostatic; proliferative; vunerary; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
 CC antiinflammatory; antihypertensive; antiallergic; antibacterial;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antitumor. The

CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilization of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AACT8457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.

CC Sequence 153 AA;

Query Match 17.3%; Score 62; DB 21; Length 153;
 Best Local Similarity 100.0%; Pred. No. 5.4e-52;
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRATPLAAPAGSLSRKKLELDNLDTERPVQKRRASGPRLPCLPLSPPTAPDRAT 60
 Db 74 MRATPLAAPAGSLSRKKLELDNLDTERPVQKRRASGPRLPCLPLSPPTAPDRAT 133

Qy 61 AV 62
 Db 134 AV 135

RESULT 12

ID ABG04413 standard; Protein; 178 AA.

AC ABG04413;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #4404.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PE 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEO INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS68600.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 20; SEQ ID No 34772; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences (I) is useful as hybridization probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 178 AA;

Query Match 10.3%; Score 37; DB 22; Length 178;
 Best Local Similarity 100.0%; Pred. No. 1.3e-27;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 RRPVQKRRASGPRLPCLPLSPPTAPDRATAVAT 64
 Db 110 RRPVQKRRASGPRLPCLPLSPPTAPDRATAVAT 146

RESULT 13

ID ABG58479 standard; Peptide; 138 AA.

AC ABG58479;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 37127.

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

PN WO200157273-A2.

PD 09-AUG-2001.

PE 30-JAN-2001; 2001WO-US00664.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488898/53.

PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -

CC Claim 27; SEQ ID No 37127; 658bp; English.
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: January 16, 2004, 10:03:56 (Search time 2407.84 Seconds

(without alignments)
10840.821 Million cell updates/sec

Title: US-09-909-474d-1_COPY_49_1122

Perfect score: 1 atgcgagccacccctctgcgc.....gagaggtctctctatgcgc 1074

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

EST:
1: em_eatba:*
2: em_eathum:*
3: em_eatlin:*
4: em_eatnu:*
5: em_eatcov:*
6: em_eatcpl:*
7: em_eatcro:*
8: em_eatc:*
9: gb_eat1:*
10: gb_eat2:*
11: gb_eatc:*
12: gb_eat3:*
13: gb_eat4:*
14: gb_eat5:*
15: em_eatfun:*
16: em_eatcom:*
17: em_g88_hum:*
18: em_g88_inv:*
19: em_g88_pln:*
20: em_g88_vtc:*
21: em_g88_fun:*
22: em_g88_mam:*
23: em_g88_mus:*
24: em_g88_pro:*
25: em_g88_rtd:*
26: em_g88_phg:*
27: em_g88_vrl:*
28: gb_g881:*
29: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	812.4	75.6	863	10	BG748142
2	802.4	74.7	1036	13	BQ217691
3	801.4	74.6	873	13	BQ216198
4	800.8	74.6	945	13	BX325345

5	795.6	74.1	902	13	BX421477
6	790.2	73.6	929	13	BQ500619
7	770.4	71.7	1201	9	AL556690
8	768.6	71.6	948	13	BX443303
9	758.2	70.6	918	13	BQ223497
10	756	70.4	914	13	BQ500176
11	744	69.3	820	13	B1256421
12	740	68.9	929	9	AL522632
13	735.4	68.6	892	14	CA95347
14	717.6	66.8	1126	13	BX405957
15	711	65.2	944	13	BX363066
16	707	65.8	1201	13	BX445734
17	705.8	65.7	1076	13	BX367265
18	703.4	65.5	838	10	BG171687
19	701.4	65.3	738	12	BM007742
20	696.6	64.9	851	12	B1909168
21	694.4	64.7	913	13	BQ433764
22	692.6	64.5	880	10	BG470123
23	688.4	64.1	851	10	BG331647
24	687.8	64.0	1053	13	BQ224844
25	687	64.0	989	10	BE792929
26	686	63.9	886	12	BM764986
27	681.2	63.4	766	12	B1253633
28	680.2	63.3	810	12	BM045379
29	679.8	63.3	748	12	B1085445
30	666.6	62.1	931	10	BG470193
31	665.2	61.9	957	13	BQ949069
32	661.4	61.6	910	10	BG397691
33	659.4	61.4	891	9	AL555267
34	655.6	61.0	716	10	BG422518
35	654.4	60.9	703	12	B1334267
36	649.8	60.5	1922	11	AK089931
37	648	60.3	648	12	BM743773
38	647.8	60.3	947	13	BQ928125
39	645	60.1	645	12	BM772759
40	641.2	59.7	878	10	BP309552
41	637.6	59.4	686	10	BG750849
42	637.4	59.3	932	13	BQ683898
43	636	59.2	648	12	BM791610
44	630.2	58.7	943	13	BQ931607
45	627.2	58.4	946	13	BQ158759

ALIGNMENTS

RESULT 1
BG748142
LOCUS
DEFINITION BG748142 863 bp mRNA EST 15-MAY-2001
mRNA sequence.
ACCESSION BG748142.1 GI:14058795
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 863)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Straube, Ph.D.
Email: g88@nci.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNU at:
http://image.llnl.gov
Plate: LNCM1676 row: 1 column: 13
High quality sequence stop: 833.

FEATURES
Source

Location/Qualifiers
1..863
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4842324"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_43"
/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library." |

BASE COUNT 135 a 300 c 264 g 164 t
ORIGIN

Query Match 75.6%; Score 812.4; DB 10; Length 863;
Best Local Similarity 98.6%; Pred. No. 4.5e-159;
Matches 830; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

```

1 ATGCGAGCCACCCCTCTGCTCTCTCTGCGGGTTCCTGTCCAGAAAGAGCGTTGAG 60
23 ATGCGAGCCACCCCTCTGCTCTCTCTGCGGGTTCCTGTCCAGAAAGAGCGTTGAG 82
61 TTGATGACAACTTGAATACGAGAGTCCCGTCAAGAAAGAGCTGAGTGGCCCGC 120
83 TTGATGACAACTTGAATACGAGAGTCCCGTCAAGAAAGAGCTGAGTGGCCCGC 142
121 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
143 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
181 GCTGTGAGCACTGCTCTCCGCTCTTGGGCGCTTATGCTCTCTGAGAGCCGAGAG 240
203 GCTGTGAGCACTGCTCTCCGCTCTTGGGCGCTTATGCTCTCTGAGAGCCGAGAG 262
241 GCGGCTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
263 GCGGCTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 322
301 GTTCAGAAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
323 GTTCAGAAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382
361 GCTCGGCACTGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
383 GCTCGGCACTGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442
421 CATGGGAGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
443 CATGGGAGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502
481 GTGCTCTTCCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
503 GTGCTCTTCCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
541 GGTGATCTTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
563 GGTGATCTTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
601 GAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
623 GAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
661 GGTGCTCCAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
683 GGTGCTCCAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
721 GAGAGCAATGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

```

```

DB 743 GCAGCGATGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
QY 761 TTCAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 803 TTCAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861
QY 841 CC 842
DB 862 CC 863

```

RESULT 2
LOCUS BQ217691
DEFINITION AGENCOURT 7570945 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6058546
5' mRNA sequence.

ACCESSION BQ217691
VERSION BQ217691.1 GI:20399091
KEYWORDS EST.

SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Buzayotai, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1036)
NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: gsa@bbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM13123 row: m column: 11
High quality sequence stop: 640.

FEATURES
Source

Location/Qualifiers
1..1036

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6058546"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_68"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies." |

BASE COUNT 173 a 365 c 301 g 196 t 1 others
ORIGIN

Query Match 74.7%; Score 802.4; DB 13; Length 1036;
Best Local Similarity 97.9%; Pred. No. 5.7e-157;
Matches 866; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

```

46 AAGAACGGTGTGAGTGTGATGACAACTTAACTACGAGCGTCCGTCAGAAAGAGCT 105
1 AAGAACGGTGTGAGTGTGATGACAACTTAACTACGAGCGTCCGTCAGAAAGAGCT 60
106 CGAAGTGGGCCCCCAGACCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 165
61 CGAAGTGGGCCCCCAGACCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
166 CGAAGTGGGCCCCCAGACCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 225
121 CGAAGTGGGCCCCCAGACCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
226 CCGAGAGAGGCGGCGGCGGCTTACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 285

```

Db	181	CCCGAGAGAGGGGGGGGGCCCTACCGGGGGCCCTGCACTGGCCCTTACAGAGGCACTGAGTATACC	240
Qy	286	TGCAGAGGTATACCCCGGTCAAGAGAGCCCTTGGCCGTGTGAGAGCCCTTACGAGCGGTGCCC	345
Db	241	TGCAGAGGTATACCCCGGTCAAGAGAGCCCTTGGCCGTGTGAGAGCCCTTACGAGCGGTGCCC	300
Qy	346	CCGCAACAGCATGTGGCTCGAGCCCACTAGAGTCTTGGCTGTATACCAGTCTCTTACAGCC	405
Db	301	CCGCAACAGCATGTGGCTCGAGCCCACTAGAGTCTTGGCTGTATACCAGTCTCTTACAGCC	360
Qy	406	TTTTTCACTCGGAGCCATGGGGACATGCACAGCCTTGTGTGGAAGCCGCGACCGTATCCCT	465
Db	361	TTTTTCACTCGGAGCCATGGGGACATGCACAGCCTTGTGTGGAAGCCGCGACCGTATCCCT	420
Qy	466	GAGCCTTGAAGCTGGCCGTCTTTCCGCGAAGTGGCACCGCCCTGGGCGCATGTATCCAG	525
Db	421	GAGCCTTGAAGCTGGCCGTCTTTCCGCGAAGTGGCACCGCCCTGGGCGCATGTATCCAG	480
Qy	526	CACGGTCTGTCTCTGCGATCTCAAGCTGTGTGCTTTGTCTTGTGCTGACCGTGAAGAG	585
Db	481	CACGGTCTGTCTCTGCGATCTCAAGCTGTGTGCTTTGTCTTGTGCTGACCGTGAAGAG	540
Qy	586	AAGAAAGCTGTGTCTTGGAGAACCTTGGAGAACTCTTGCCTGTGTACTTGGGCGCAGATATTC	645
Db	541	AAGAAAGCTGTGTCTTGGAGAACCTTGGAGAACTCTTGCCTGTGTACTTGGGCGCAGATATTC	600
Qy	646	CTGTGGGACAAAGCAGCGGTCCAGACCTTACGTGGAGACCTGAGATACTAGCTCACGGGCC	705
Db	601	CTGTGGGACAAAGCAGCGGTCCAGACCTTACGTGGAGACCTGAGATACTAGCTCACGGGCC	660
Qy	706	TCATATCTCGGGCAGAGCCGAGTCTTGGAGCCTTGGCGGTGTGCGCTCTTCCACATGCT-	764
Db	661	TCATATCTCGGGCAGAGCCGAGTCTTGGAGCCTTGGCGGTGTGCGCTCTTCCACATGCTG	720
Qy	765	GCGCGGCACTAACCCCTTCCAGGAATCGGAGGCTGTCCCTGCTTTGGCAAGATACC--GC	822
Db	721	GCGCGGCACTAACCCCTTCCAGGAATCGGAGGCTGTCCCTGCTTTGGCAAGATACC--GC	780
Qy	823	CGCGGGGCGCTACGCTTTCCTTGCAGAGGCTCTTCCGAGCCCTTG--CCGCTGTCTGTTCGCTG	881
Db	781	CGCGGGGCGCTACGCTTTCCTTGCAGAGGCTCTTCCGAGCCCTTG--CCGCTGTCTGTTCGCTG	840
Qy	882	CC-TCTTTCGTGGGA--CCGACGTGAACGCTCAACGCCACAG	923
Db	841	CTTCTCTTCTTCCGAGAGCCGCTTGAACGCTCAACGCCACAG	885
RESULT 3			
BQ216198			
LOCUS	BQ216198	873 bp	mRNA linear EST 02-MAY-2002
DEFINITION	AGENCOURT_7574831 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059905		
ACCESSION	BQ216198		
VERSION	BQ216198.1		
KEYWORDS	GI:20397598		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 873)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: gsa@bbs-rcmail.nih.gov		
	Tissue Procurement: DCTD/DTF/Gazdar		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNMU)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNMU at:		
	http://image.llnl.gov		
	Plate: LLM13327		
	row: 6 column: 02		

High quality sequence stop: 600.									
FEATURES		Location/Qualifiers							
Source		1..873							
		/organism="Homo sapiens"							
		/mol_type="mRNA"							
		/db_xref="taxon:9606"							
		/clone="IMAGE:605905"							
		/issue_type="large cell carcinoma"							
		/lab_host="DH10B (phage-resistant)"							
		/clone_1lb="NH MGc_68"							
		/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."							
BASE COUNT		144 a	305 c	264 g	160 t				
ORIGIN									
Query Match	74.6%; Score 801.4; DB 13; Length 873;								
Best Local Similarity	99.1%; Pred. No. 8.8e-157;								
Matches 816; Conservative	0; Mismatches 6; Indels 1; Gaps 1								
QY	1 ATGCGAGCCACCCCTCTGCTGCTCTCTCGGGGTTCCCTGTCAGAGAGAGCGTTGGAG	60							
DB	51 ATGCGAGCCACCCCTCTGCTGCTCTCTCGGGGTTCCCTGTCAGAGAGAGCGTTGGAG	110							
QY	61 TTGATATCAACTTATAGATACCGAGGTCGCGTCCAGAAACGAGCTCGAAGTGGGCGCAG	120							
DB	111 TTGATATCAACTTATAGATACCGAGGTCGCGTCCAGAAACGAGCTCGAAGTGGGCGCAG	170							
QY	121 CCCAGACTGCCCCCTGCTTGGCCCTGAGCCCACTACTGCTCAGATCTGACAACT	180							
DB	171 CCCAGACTGCCCCCTGCTTGGCCCTGAGCCCACTACTGCTCAGATCTGACAACT	230							
QY	181 GCTGTGGCCACTGCTCTCCCGTCTTGGGCGCTATGTCTCTCTGAGAGCCCGAGAGAGGCGG	240							
DB	231 GCTGTGGCCACTGCTCTCCCGTCTTGGGCGCTATGTCTCTCTGAGAGCCCGAGAGAGGCGG	290							
QY	241 CGGGCTTACCGGGCCCTGACCTGCGCTTACAGAGCACTGAGTATACCTGCAAGGTATACCC	300							
DB	291 CGGGCTTACCGGGCCCTGACCTGCGCTTACAGAGCACTGAGTATACCTGCAAGGTATACCC	350							
QY	301 GTCCAGAGAGCCCTGAGCGGTGCTGAGAGCTTACCGGCGGCTGCCCCGACAGAGATGTG	360							
DB	351 GTCCAGAGAGCCCTGAGCGGTGCTGAGAGCTTACCGGCGGCTGCCCCGACAGAGATGTG	410							
QY	361 GCTCGGCCCCCATGAGGTCTGTGGCTGTATCCAGAGCTCTTACGCGCTTTTCACTGAGAC	420							
DB	411 GCTCGGCCCCCATGAGGTCTGTGGCTGTATCCAGAGCTCTTACGCGCTTTTCACTGAGAC	470							
QY	421 CATGGGAGCATGACAGCGCTGTGTGCGAAGCGGACCGATATCCCTGAGACCTGAGGCTCC	480							
DB	471 CATGGGAGCATGACAGCGCTGTGTGCGAAGCGGACCGATATCCCTGAGACCTGAGGCTCC	530							
QY	481 GTGCTCTTCCGCGCATGAGCCACCGGCTTGGGCGCATGTCAACAGACAGGTCTGTGCTTG	540							
DB	531 GTGCTCTTCCGCGCATGAGCCACCGGCTTGGGCGCATGTCAACAGACAGGTCTGTGCTTG	590							
QY	541 CGTATCTCAAGCTGTGTGCGCTTTGTCTTGTGCTGACCGGTGAGAGAGAAAGCTGTGTG	600							
DB	591 CGTATCTCAAGCTGTGTGCGCTTTGTCTTGTGCTGACCGGTGAGAGAGAAAGCTGTGTG	650							
QY	601 GAGAACTGAGAGACTCTCTGTGTCTGATCTGAGGCGAAGTAAATTCCTGTGTGGAACAAGAC	660							
DB	651 GAGAACTGAGAGACTCTCTGTGTCTGATCTGAGGCGAAGTAAATTCCTGTGTGGAACAAGAC	710							
QY	661 GCGAGCCGAGCTTACGTGTGGGACCTGAGATTAATCAAGCTCAAGGGCTTAATCTCGGGCAAG	720							
DB	711 GCGAGCCGAGCTTACGTGTGGGACCTGAGATTAATCAAGCTCAAGGGCTTAATCTCGGGCAAG	770							
QY	721 GCAAGCCGATGCTT-GAGAGCTTGGGCGTGTGCGCTTTTCACCAATGTGAGCGGCACTACCC	779							
DB	771 GCAAGCCGATGCTTGGGCGCTTGGGCGTGTGCGCTTTTCACCAATGTGAGCGGCACTACCC	830							

QY 780 CTTCCAGAGACTCGAGAGCTGTCTCTCTTCGCAAGATCCGC 822
DB 831 CTTCCAGAGACTCGAGAGCTGTCTCTCTTCGCAAGATCCGC 873

RESULT 4
LOCUS BX325345 945 bp mRNA linear EST 02-MAY-2003
DEFINITION BX325345 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
ACCESSION Homo sapiens cDNA clone CS0DL004YF04 5-PRIME, mRNA sequence.
VERSION BX325345.1 GI:30338449
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 945)
AUTHORS Li, W.-B., Gruber, C., Jesssee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10000.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL004YF04&cluster=10000.f. Contact :
Peng Liang Email: liliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL004YF04P1.

FEATURES
Source Location/Qualifiers
1..945
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL004YF04"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_1lb="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and BclI V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 163 a 313 c 282 g 176 t
ORIGIN

Query Match 74.6%; Score 800.8; DB 13; Length 945;
Best Local Similarity 97.7%; Pred. No. 1.2e-156;
Matches 821; Conservative 10; Mismatches 7; Indels 2; Gaps 2;

QY 1 ATGGAGGACACCCCTGTGCTGTCTGTGCGGGTTCCTGTCAGGAAGAGGCTTGAG 60
DB 107 ATGGAGGACACCCCTGTGCTGTCTGTGCGGGTTCCTGTCAGGAAGAGGCTTGAG 166
QY 61 TTG-GATGACAACTTATGATACCGAGGCTCCGCTCCAGAAACGAGCTCGAAGTGGGCCCA 119
DB 167 TTGATATGACAACTTATGATACCGAGGCTCCGCTCCAGAAACGAGCTCGAAGTGGGCCCA 226
QY 120 GCCCAGACTGCCCCCTGTGCTGTGCTGTGAGCCCACTTACTGCTCCGATGTCGCAAC 179
DB 227 GCCCAGACTGCCCCCTGTGCTGTGCTGTGAGCCCACTTACTGCTCCGATGTCGCAAC 286
QY 180 TGTCTGTGACACTGCTCCGCTCTTGTGGGCTATGCTCTCTGAGAGCCGAGAGGGCGG 239
DB 287 TGTCTGTGACACTGCTCCGCTCTTGTGGGCTATGCTCTCTGAGAGCCGAGAGGGCGG 346
QY 240 GCGGGGCTTACCGGGGCTTGTGACTGCTTACGAGGCACTGAGTATTAAGTGTGAGTACC 299
DB 347 GCGGGGCTTACCGGGGCTTGTGACTGCTTACGAGGCACTGAGTATTAAGTGTGAGTACC 406

QY 300 GTTCCAGAGAGCTGTGCTGTGCTGTGAGAGCCCTTACGCGCGGCTGCCCGCACAGATGT 359
DB 407 GTTCCAGAGAGAGCTGTGCTGTGCTGTGAGAGCCCTTACGCGCGGCTGCCCGCACAGATGT 466
QY 360 GGCTCCGGCCCACTGAGAGCTGTGCTGTGAGAGCCCTTACGCGCGGCTGCCCGCACAGATGT 419
DB 467 GGCTCCGGCCCACTGAGAGCTGTGCTGTGAGAGCCCTTACGCGCGGCTGCCCGCACAGATGT 526
QY 420 CCATGAGGACATGACACAGCTGTGCTGTGAGAGCCCGGACCGGATATCCCTGAGCTGAGCTGC 479
DB 527 CCATGAGGACATGACACAGCTGTGCTGTGAGAGCCCGGACCGGATATCCCTGAGCTGAGCTGC 585
QY 480 CGTGTCTTTCGCGCAGATGCGCACCGGCTGTGAGAGCTGTGACACGAGTGTGCTCT 539
DB 586 CGTGTCTTTCGCGCAGATGCGCACCGGCTGTGAGAGCTGTGACACGAGTGTGCTCT 645
QY 540 GCGTATCTCAAGCTGTGCTGTGCTGTGAGAGCTGTGAGAGAGAGAGTGTGCT 599
DB 646 GCGTATCTCAAGCTGTGCTGTGCTGTGAGAGCTGTGAGAGAGAGAGTGTGCT 705
QY 600 GGAGAACCTTGAAGAGCTGTGCTGTGAGAGCTGTGAGAGAGAGAGTGTGCTGTGAGAGCA 659
DB 706 GGAGAACCTTGAAGAGCTGTGCTGTGAGAGCTGTGAGAGAGAGAGTGTGCTGTGAGAGCA 765
QY 660 CGCGTGTGCGCAGCTTGTGAGAGCTGTGAGAGCTGTGAGAGAGAGAGTGTGCTGTGAGAGCA 719
DB 766 CGCGTGTGCGCAGCTTGTGAGAGCTGTGAGAGCTGTGAGAGAGAGAGTGTGCTGTGAGAGCA 825
QY 720 GCGAGCGGATGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGAGAGAGTGTGAGAGCA 779
DB 826 GCGAGCGGATGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGAGAGAGTGTGAGAGCA 885
QY 780 CTTCCAGAGACTCGAGAGCTGTCTCTCTTCGCAAGATCCGC 839
DB 886 CTTCCAGAGACTCGAGAGCTGTCTCTCTTCGCAAGATCCGC 945

RESULT 5
LOCUS BX421477 902 bp mRNA linear EST 15-MAY-2003
DEFINITION BX421477 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
clone CS0DG007YG03 5-PRIME, mRNA sequence.
ACCESSION BX421477.1 GI:30764090
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 902)
AUTHORS Li, W.-B., Gruber, C., Jesssee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10000.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DG007YG03&cluster=10000.f. Contact :
Peng Liang Email: liliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DG007YG03P1.

FEATURES
Source Location/Qualifiers
1..902
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG007YG03"
/issue_type="B CELLS (RAMOS CELL LINE)"

DEFINITION	BK443303 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA clone CS00G007YP05 5-PRIME, mRNA sequence.
ACCESSION	BK443303
VERSION	BK443303.1
KEYWORDS	EST.
SOURCE	GI:31018564
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li, W.B., Gruber, C., Jeehee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10000.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DG007CC030PLC1cluster=10000.f. Contact : Peng Liang Email : fliang@life.techn.com URL : Corporation 1600 http://fulllength.invitrogen.com/ Invitrogen 1600 Paradey Avenue Genoscope sequence ID : CS0DG007CC030PL. Location/Qualifiers
FEATURES	1..948
source	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DG007YP05" /tissue_type="B CELLS (RAMOS CELL LINE)" /cell_line="RAMOS CELL LINE" /clone_idb="Homo sapiens B CELLS (RAMOS CELL LINE)" /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT	152 a 326 c 296 g 169 t 5 others
ORIGIN	
Query Match	71.6%; Score 768.6; DB 13; Length 948;
Beet Local Similarity	99.4%; Pred. No. 6e-150; 1; Indels 2; Gaps 2;
Matches 790; Conservative	2; Mismatches 1; Indels 2; Gaps 2;
QY	1 ATGCAGACCAACCCCTGCTGCTCTGCTCTGCGAGTTCCTCTGCAGAGAAACGGTTGAG 60
DB	156 ATGGAGCCACCCCTCTGGCTGCTCTGCGAGTTCCTCTGCAGAGAAACGGTTGAG 215
QY	61 TTGGATGCAACTTAGATACCGAGGCTCCCTCCAGAAACGAGCTCGAAGTGGCCCCAG 120
DB	216 TTGGATGCAACTTAGATACCGAGGCTCCCTCCAGAAACGAGCTCGAAGTGGCCCCAG 275
QY	121 CCAGACAGTCCGCCCTGCTGTTGGCCCTGAGCCCACTATGCTCCAGATGTGCACT 180
DB	276 CCAGACAGTCCGCCCTGCTGTTGGCCCTGAGCCCACTATGCTCCAGATGTGCACT 335
QY	181 GCTGTGGCACTGCTCCCGTCTTGAGGCCATATGTCTCTCTGAGACCCGAGAGAGCGCG 240
DB	336 GCTGTGGCACTGCTCCCGTCTTGAGGCCATATGTCTCTCTGAGACCCGAGAGAGCGCG 395
QY	241 CGGAGCTACCGGGCCCTGCACTGGCCCTACAGGACTAGATTAACCTGCAAGTGTACCC 300
DB	396 CGGAGCTACCGGGCCCTGCACTGGCCCTACAGGACTAGATTAACCTGCAAGTGTACCC 454
QY	301 GTCCAGAGAACCTTGCGCTGCTGAGAGCCCTACGCGGAGCTGCCCCGACAAAGATATG 360
DB	455 GTCCAGAGAACCTTGCGCTGCTGAGAGCCCTACGCGGAGCTGCCCCGACAAAGATATG 514
QY	361 GCTGGGCCCACTGAGAGTCTGGGCTGATACCAAGCTCTCTTAGAGCTTTTTCATCTGAGAC 420
DB	515 GCTGGGCCCACTGAGAGTCTGGGCTGATACCAAGCTCTCTTAGAGCTTTTTCATCTGAGAC 574

QY	421	CATGGGGACATGACACAGCTGTGTGCGCTTTGCTTTCGCTGACCGGTATCCCTGAGACCTGAGGCTGCGC	480		
Db	575	CATGGGGACATGACACAGCTGTGTGCGCTTTGCTTTCGCTGACCGGTATCCCTGAGACCTGAGGCTGCGC	633		
QY	481	GTGCTCTTCCGACAGATGAGCCACCGCCCTGTGGCGGACTGTCAACGACACGATCTGTGCTGTG	540		
Db	634	GTGCTCTTCCGACAGATGAGCCACCGCCCTGTGGCGGACTGTCAACGACACGATCTGTGCTGTG	693		
QY	541	CGTATCTCAAGCTGTGTGTGCTTTGCTTTCGCTGACCGGTATCCCTGAGACCTGAGTGTG	600		
Db	694	CGTATCTCAAGCTGTGTGTGCTTTGCTTTCGCTGACCGGTATCCCTGAGACCTGAGTGTG	753		
QY	601	GAGAACCTGAGAGACTCCTGTGGCTGTGACCTGGGCGAGATGATTCCTGTGGAGACAAGAC	660		
Db	754	GAGAACCTGAGAGACTCCTGTGGCTGTGACCTGGGCGAGATGATTCCTGTGGAGACAAGAC	813		
QY	661	GGGTGCCAGGCTACGTGGAGACTGAGATACCTGACGTACAGGAGCTGATACCTCGGCAAG	720		
Db	814	GGGTGCCAGGCTACGTGGAGACTGAGATACCTGACGTACAGGAGCTGATACCTCGGCAAG	873		
QY	721	GCAGCCGATGCTGTGAGCTGTGGCGGCTTTACCATGTGGCCGACCTACCC	780		
Db	874	GCAGCCGATGCTGTGAGCTGTGGCGGCTTTACCATGTGGCCGACCTACCC	933		
QY	781	TTCCAGGACTGGAG	795		
Db	934	TTCCAGGACTGGAG	948		
RESULT 9	BQ223497	918 bp	mRNA	linear	EST 02-MAY-2002
LOCUS	BQ223497				
DEFINITION	AGENCOURT_7576910 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6058986				
ACCESSION	BQ223497				
VERSION	BQ223497.1	GI:20404897			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukariyola, Metaxoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 918)				
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
UNPUBLISHED	Contact: Robert Strausberg, Ph.D.				
	Email: csapbs-remail.nih.gov				
	Tissue Procurement: DCTD/DTF/Gazdar				
	cDNA Library Preparation: Life Technologies, Inc.				
	DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.lnl.gov				
	Plate: LHAM3324	row: 0	column: 19		
	High quality sequence, stop: 684.				
FEATURES	Location/Qualifiers				
source	1..918				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:6058986"				
	/tissue_type="large cell carcinoma"				
	/lab_host="DH10B (phage-resistant)"				
	/clone_id="NIH MGC 68"				
	/notes="Organ: lung; Vector: pCMV-SPORT6, site 1: NotI;				
	site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.				
	Average insert size 1.8 kb. Library constructed by Life				
	Technologies."				
BASE COUNT	144 a	321 c	285 g	168 t	
ORIGIN					

Query Match 70.6%; Score 758.2; DB 13; Length 918;
 Best Local Similarity 98.5%; Pred. No. 8.7e-148;
 Matches 807; Conservative 0; Mismatches 8; Indels 4; Gaps 4;

```

OY 1 ATGAGAGCACCCCTCTGGCTGCTCTCTGGGGGTTCTCTGACGAGAAAGAGCGTTGAG 60
DB 100 ATGAGAGCACCCCTCTGGCTGCTCTCTGGGGTTCCTGTCAGAGAAAGAGCGTTGAG 159
OY 61 TTGATGACAACTTAGATACCGAGCGTCCCTGTCAGAAAAGAGCTGAAAGTGGCCCCAG 120
DB 160 TTGATGACAACTTAGATACCGAGCGTCCCTGTCAGAAAAGAGCTGAAAGTGGCCCCAG 219
OY 121 CCCAGACTGCCCCCTGCTGTTGGCCCTGAGCCCACTGCTGCTGAGATGTGCAACT 180
DB 220 CCCAGACTGCCCCCTGCTGTTGGCCCTGAGCCCACTGCTGCTGAGATGTGCAACT 279
OY 181 GCTGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 280 GCTGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 339
OY 241 CGGAGCTACCGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 340 CGGAGCTACCGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 399
OY 301 GTCCAGAAAGCCCTGCGGCTGCTGAGCCCTGAGCCGCTGAGCCGCTGAGCCGCTGAG 360
DB 400 GTCCAGAAAGCCCTGCGGCTGCTGAGCCCTGAGCCGCTGAGCCGCTGAGCCGCTGAG 459
OY 361 GCTCGGCCCACTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 460 GCTCGGCCCACTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
OY 421 CATGGGGAATGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 520 CATGGGGAATGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
OY 481 GTGCTCTTCCGCGAGATGAGCCAGCCGCTGAGCCGCTGAGCCGCTGAGCCGCTGAG 540
DB 580 GTGCTCTTCCGCGAGATGAGCCAGCCGCTGAGCCGCTGAGCCGCTGAGCCGCTGAG 639
OY 541 CGTGAATCTCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 640 CGTGAATCTCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 699
OY 601 GAGAACTTGAAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 700 GAGAACTTGAAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759
OY 661 GCGTCCCAAGCTTACGCTGAGACTGAGTACTGAGTACTGAGTACTGAGTACTGAGTACTG 720
DB 760 GCGTCCCAAGCTTACGCTGAGACTGAGTACTGAGTACTGAGTACTGAGTACTGAGTACTG 819
OY 721 GCAAGCGATGTCTGAGAGCTGAGAGC-TGGCGCTCTTACCAATGCT-GGCGGCGCACTAC 778
DB 820 GCAAGCGATGTCTGAGAGCTGAGAGC-TGGCGCTCTTACCAATGCT-GGCGGCGCACTAC 879
OY 779 CC-TTCCAGAGACTC-GAGAGCTGTCTGCTCTTCCGCA 815
DB 880 CTTTCCAGAGACTC-GAGAGCTGTCTGCTCTTCCGCA 918

```

RESULT 10
 BUS00176 914 bp mRNA 1linear EST 12-SEP-2002
 LOCUS AGSCOURT_7859610 NIH_MGC_64 Homo sapiens cDNA clone IMAGE:6108987
 DEFINITION 5', mRNA sequence.

ACCESSION BUS00176
 VERSION BUS00176.1 GI:22800682
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 914)
 AUTHORS NIH-MGC http://mgi.nhl.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bme.nhl.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: NIMH/LOG
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 P1: LCM2354 row: c column: 04
 High quality sequence scop: 682.
 Location/Qualifiers

FEATURES
 source
 1. 914
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /lab_host="DH10B (phage-resistant)"
 /clone="NIH MGC 64"
 /note="Vector: pOTB7a; Site 1: CeuI; Site 2: ScaI; This
 library is a size selection of NIH MGC 35, from 3.0-4.5
 kb. Size selection done at the National Institute of
 Mental Health, NIH. Note: this is a NIH MGC library."

BASE COUNT 147 a 323 c 274 g 169 t
 ORIGIN

Query Match 70.4%; Score 756; DB 13; Length 914;
 Best Local Similarity 96.1%; Pred. No. 2.5e-147;
 Matches 808; Conservative 0; Mismatches 26; Indels 7; Gaps 3;

```

OY 49 AAGCGTGTGAGTTGATGACAACTTAGATACCGAGCGTCCGTCAGAAAGAGCTGCA 108
DB 1 AAGCGTGTGAGTTGATGACAACTTAGATACCGAGCGTCCGTCAGAAAGAGCTGCA 60
OY 109 AGTGGGCCCCAGCCAGACTGCCCCCTGCTGTTGGCCCTGAGCCCACTACTGCTCA 168
DB 61 AGTGGGCCCCAGCCAGACTGCCCCCTGCTGTTGGCCCTGAGGCCCACTACTGCTCA 120
OY 169 GATCGTGAACCTGCTGAGCCCACTGCTCCGCTTGGGAGCCCTATATGCTCTCTGAGGCC 228
DB 121 GATCGTGAACCTGCTGAGCCCACTGCTCCGCTTGGGAGCCCTATATGCTCTCTGAGGCC 180
OY 229 GAGAGGCGGCGGAGCTTACCGGCGCTGCACTGCGCTTACAGGCACTGAGTATCTGCG 288
DB 181 GAGAGGCGGCGGAGCTTACCGGCGCTGCACTGCGCTTACAGGCACTGAGTATCTGCG 240
OY 289 AAGGTGTACCCCGTCCAGAAAGCCCTGCGCGTGTGAGAGCCCTTACCGCGAGTCCCGG 348
DB 241 AAGGTGTACCCCGTCCAGAAAGCCCTGCGCGTGTGAGAGCCCTTACCGCGAGTCCCGG 300
OY 349 CACAAGCATGTGAGCTCGGAGCACTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 408
DB 301 CACAAGCATGTGAGCTCGGAGCACTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
OY 409 TTCACTCGAAGCCATGAGGAGCAATGCAAGCTGTGTGAGAGCCGCAACCGTATCCCTGAG 468
DB 361 TTCACTCGAAGCCATGAGGAGCAATGCAAGCTGTGTGAGAGCCGCAACCGTATCCCTGAG 420
OY 469 CTTGAGGCTGCGCTGCTTTCGCGAGATGAGCAACCGGCTGAGGAGCACTGTACCAAGCAC 528
DB 421 CTTGAGGCTGCGCTGCTTTCGCGAGATGAGCAACCGGCTGAGGAGCACTGTACCAAGCAC 480
OY 529 GGTCTGTGCTCTGCGTATCTCAAGCTGTGTGCTTCTTCTGCTGAGCCGCTGAGAGAG 588
DB 481 GGTCTGTGCTCTGCGTATCTCAAGCTGTGTGCTTCTTCTGCTGAGCCGCTGAGAGAG 540
OY 589 AAGCTGTGCTGAGAGACTTGAAGAGACTCTGCGTGTGAGTCTGAGGCGCAAGTATTCCTG 648
DB 541 AAGCTGTGCTGAGAGACTTGAAGAGACTCTGCGTGTGAGTCTGAGGCGCAAGTATTCCTG 600

```


649 TGGGCAAGCAGCGCTGCGCAGCTTACGCTGAGGACCTGAGATATCTAGCTACCGGCTCA 708
DB 601 TGGGCAAGCAGCGCTGCGCAGCTTACGCTGAGGACCTGAGATATCTAGCTACCGGCTCA 660
609 TACTCGGCAAGCAGCGCTGCGCAGCTTACGCTGAGGACCTGAGATATCTAGCTACCGGCTCA 767
DB 661 TACTCGGCAAGCAGCGCTGCGCAGCTTACGCTGAGGACCTGAGATATCTAGCTACCGGCTCA 720
678 CGGCACTACCTGCGCAGCTTACGCTGAGGACCTGAGATATCTAGCTACCGGCTCA 824
DB 721 CGGCACTACCTGCGCAGCTTACGCTGAGGACCTGAGATATCTAGCTACCGGCTCA 780
625 CGGCACTACCTGCGCAGCTTACGCTGAGGACCTGAGATATCTAGCTACCGGCTCA 881
DB 781 CGGCACTACCTGCGCAGCTTACGCTGAGGACCTGAGATATCTAGCTACCGGCTCA 840
682 C 882
DB 841 C 841

RESULT 11
BI256421 820 bp mRNA linear EST 17-JUL-2001
LOCUS 602974439P1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5113905 5',
DEFINITION mRNA sequence.
ACCESSION BI256421.1 GI:14810812
VERSION BI256421.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 820)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Straube, Ph.D.
Email: cgarbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLND at:
http://image.llnl.gov
Plate: L1AM11276 row: m column: 10
High quality sequence spot: 820.
Location/Qualifiers
1..820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5113905"
/label_type="cervical carcinoma cell line"
/lab_host="DH108"
/clone_id="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-Sport6; Site: 1; NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 133 a 285 c 245 g 157 t
ORIGIN

Query Match 69.3%; Score 744; DB 12; Length 820;
Best Local Similarity 98.7%; Pred. No. 7.7e-145;
Matches 813; Conservative 0; Mismatches 5; Indels 6; Gaps 6;
43 AGGAGAGAGGCTTGGATGATGACACTTATGATCCGAGCGCTCCGTCAGAAAGCA 102
DB 1 AGGAGAGAGGCTTGGATGATGACACTTATGATCCGAGCGCTCCGTCAGAAAGCA 60

103 GCTGAGATGAGGCCCCAGACCCAGATGCGCCCCCTGCTGTTGAGCCCTGAGCCACTACT 162
DB 61 GCTGAGATGAGGCCCCAGACCCAGATGCGCCCCCTGCTGTTGAGCCCTGAGCCACTACT 120
163 GCTGAGATGAGGCCCCAGATGCGCCCCCTGCTGTTGAGCCCTGAGCCACTACT 222
DB 121 GCTGAGATGAGGCCCCAGATGCGCCCCCTGCTGTTGAGCCCTGAGCCACTACT 180
223 GAGCCGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 282
DB 181 GAGCCGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
283 ACTGCAAGATGATACCCGCTCAGAGAGCCGCTGAGGCGGCTGAGGCGGCGGCGGCGG 342
DB 241 ACTGCAAGATGATACCCGCTCAGAGAGCCGCTGAGGCGGCTGAGGCGGCGGCGGCGG 300
343 CCCCCG-CACAGCATGATGAGCTGCGCCCACTGAGGCTGCTGAGTACCACTCTCTA 401
DB 301 CCCCCG-CACAGCATGATGAGCTGCGCCCACTGAGGCTGCTGAGTACCACTCTCTA 360
402 CGCTTTTTCATCTGGAACCATGAGGACATGACAGCTGCTGAGTACCACTCTCTA 461
DB 361 CGCTTTTTCATCTGGAACCATGAGGACATGACAGCTGCTGAGTACCACTCTCTA 420
462 CCTGAGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 521
DB 421 CCTGAGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
522 CCAGCAGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 581
DB 481 CCAGCAGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
582 GAGGAGAGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 641
DB 541 GAGGAGAGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
642 TTCCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 701
DB 601 TTCCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
702 GAGCTCATCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 760
DB 661 GAGCTCATCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
761 TGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 820
DB 721 TGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 779
821 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 884
DB 780 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 820

RESULT 12
AL522632 929 bp mRNA linear EST 22-MAY-2003
LOCUS AL522632 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0BD009Y003 5-PRIME, mRNA sequence.
ACCESSION AL522632
VERSION AL522632.2 GI:31040900
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS Li, W. B., Gruber, C., Jassie, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 13, 2001 this sequence version replaced gi:12786125.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

Email: beqref@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10000.f For
more information about this cluster, see
http://www.genoscope.cns.fr/

cg1-bin/cluster.cgi?seq=CS0DB009AH020P1&cluster=10000.f. Contact :
Feng Liang Email: fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Parade Avenue Genoscope sequence ID : CS0DB009AH020P1.

FEATURES

source

1..929
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB009Y003"
/feature_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
284 g 313 c 284 g 169 c 5 others

BASE COUNT

158 a 313 c 284 g 169 t

ORIGIN

Query Match 68.9%; Score 740; DB 9; Length 929;

Best Local Similarity 99.5%; Pred. No. 5,4e-144;

Matches 762; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

1 ATGGCAGCCACCCCTGTGCTGCTGCGGGGTTCCCTGCTCCAGAAAGACGGTGGAG 60
166 ATGCAGACCACTTGTGCTGCTGCTGCGGGTTCCTTCCAGAAAGACGGTGGAG 225
61 TTGATGACAACTTGTGATACCGAGCGTCCGTTGAGAAAGAGCTGAAAGTGGCCAG 120
226 TTGATGACAACTTGTGATACCGAGCGTCCGTTGAGAAAGAGCTGAAAGTGGCCAG 285
121 CCCAGATGCCCCCTGCTGCTGCTGCGGGGCTTCCCTGCTCCAGAAAGACGGTGGAG 180
286 CCCAGATGCCCCCTGCTGCTGCTGCGGGGCTTCCCTGCTCCAGAAAGACGGTGGAG 345
181 GCTGTGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
346 GCTGTGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 405
241 CGGCGCTTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
406 CGGCGCTTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464
301 GTCGAGGAAGCCCTGCGCTGCGTGAAGCCCTTACGCGCGGCTGCCCGGCAAGAGCAATG 360
465 GTCGAGGAAGCCCTGCGCTGCGTGAAGCCCTTACGCGCGGCTGCCCGGCAAGAGCAATG 524
361 GCTCGGCGCACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
525 GCTCGGCGCACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 584
421 CATGGGAGACATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
585 CATGGGAGACATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
481 GTGCTCTTCCGCGAGATGGCAACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 540
644 GTGCTCTTCCGCGAGATGGCAACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 703
541 CGTGATCTCAAGCTGT 600
704 CGTGATCTCAAGCTGT 763
601 GAGAACTTGAAGACTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
764 GAGAACTTGAAGACTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823
661 GCGTGGCCAGCTAAGTGGGAGCTGAGATCTAGACTCAGGCGCTCATCTCGGCGAAG 720

|||||
Db 824 GCGTCCCAAGCTTACGAGCTGAGACCTGAGATCTCACTCAAGGCGCTTATCTCGGCGAAG 883
Oy 721 GCAGCGGATGCTGAGAGCTGCGCGCTGCGCTTTCACCATGCTGG 766
Db 884 GCAGCGGATGCTGAGAGCTGCGCGCTGCGCGCTTTCACCATGCTGG 929

RESULT 13

CA495347

LOCUS

AGENCOURT_10953584

DEFINITION

MRCN sequence.

ACCESSION

CA495347

VERSION

CA495347.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 892)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapdb-remail.nih.gov

Tissue Procurement: Kristi A. Egland, Ira Pastan

cDNA Library Preparation: Invitrogen Corp

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLM14286

row: 1

column: 21

High quality sequence stop: 653.

Location/Qualifiers

1..892

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGR:6785470"

/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HMEL

LNCaP"

/lab_host="BMDH10B"

/clone_lib="MAPCL"

/note="Vector: PCMV-SPORT6; Site 1: BclRV; Site 2: Not I;

Subtracted with brain, liver, lung, kidney and muscle.

Directionally cloned. Priming method: oligo-dT. Average

insert size: 1800 bp. Library amplification: 26,000 fold.

Kristi A. Egland, James J. Vincent, Robert Strausberg,

Bungrook Lee & Ira Pastan: Discovery of new breast

cancer genes encoding membrane and secreted proteins.

Manuscript submitted."

BASE COUNT

155 a 309 c 267 g 161 t

ORIGIN

Query Match 68.6%; Score 736.4; DB 14; Length 892;
Best Local Similarity 97.5%; Pred. No. 3e-143;
Matches 760; Conservative 0; Mismatches 16; Indels 4; Gaps 3;

215 TCCTCTGAGAGCCGAGAGAGGCGGCGGCGCTACCGGCGCTGCACTGCGCTCAAGGCA 274
1 TCCTCTGAGAGCCGAGAGAGGCGGCGGCGCTACCGGCGCTGCACTGCGCTCAAGGCA 60
Oy 275 CTGAGATATCTGCAAGGTGTATCCCGTCCAGAGAGCCCTGCGCTGCTGAGAGCCCTTACG 334
Db 61 CTGAGATATCTGCAAGGTGTATCCCGTCCAGAGAGCCCTGCGCTGCTGAGAGCCCTTACG 120
Oy 335 CGCGGCTGCGCGCGCAAGCATGTGGCTGCGCGCCCACTGAGGTCTGCGTGTATACCAAGC 394
Db 121 CGCGGCTGCGCGCGCAAGCATGTGGCTGCGCGCCCACTGAGGTCTGCGTGTATACCAAGC 180
Oy 395 TCCTCTGAGAGCCCTTTCCTGCACTGCGAGCCATGTGGGAGCATGCAAGCGCTGCTGAGAGCGCGC 454

Db	181	TCCTTAAAGCCTTTTTCACCTGGACCCATGGGGACATGACAGCTGGTGGCAACCCGCC	240
Oy	455	ACCGTATCCCTGAGCCCTGAGGCTGGCCGTGCTCTTCCGCGAATGACACCGCCCTGGCGC	514
Db	241	ACCGTATCCCTGAGCCCTGAGGCTGGCCGTGCTCTTCCGCGAATGACACCGCCCTGGCGC	300
Oy	515	ACTGTCACAGACGGTCTGTCGTGCTGGCGATCTGACGCTGTGTGGCTTTGTCTTGGCTG	574
Db	301	ACTGTCACAGACGGTCTGTCGTGCTGGCGATCTGACGCTGTGTGGCTTTGTCTTGGCTG	360
Oy	575	ACCGTGAAGAGAAAGCTGTGTCTGAGAAACCTGAGAGACTCTGCGTGTGACTG	634
Db	361	ACCGTGAAGAGAAAGCTGTGTCTGAGAAACCTGAGAGACTCTGCGTGTGACTG	420
Oy	635	CAGATGATTTCTCTGTGGGACAAAGCAACGCTGGCCAGCTTACCTGTGGACCTGATATCTCA	694
Db	421	CAGATGATTTCTCTGTGGGACAAAGCAACGCTGGCCAGCTTACCTGTGGACCTGATATCTCA	480
Oy	695	GCTCACGGGGCTCATACTCTGGGCAAGCAGCGATGTCTGAGACCTGGGGCTGGCGCTCT	754
Db	481	GCTCACGGGGCTCATACTCTGGGCAAGCAGCGATGTCTGAGACCTGGGGCTGGCGCTCT	540
Oy	755	TCACCATGTGGCCGGGCCACTACCCCTTTCAGAGACTCGAGGCTGTCTGTCTTTGGGCA	814
Db	541	TCACCATGTGGCCGGGCCACTACCCCTTTCAGAGACTCGAGGCTGTCTGTCTTTGGGCA	600
Oy	815	AGATCCGCGCGCGGGGCTTACGCGCTTGCTGCGAGGCTCTCGGGCCCTGCGCGCTGTGG	874
Db	601	AGATCCGCGCGCGGGGCTTACGCGCTTGCTGCGAGGCTCTCGGGCCCTGCGCGCTGTGG	660
Oy	875	TTGCGTGCCTCTCTTGTGTGGGAGCCAGCTGAAACGAGCTCAACAGCCACAGGCAATCTCTGC	934
Db	661	TTGCGTGCCTCTCTTGTGTGGGAGCCAGCTGAAACGAGCTCAACAGCCACAGGCAATCTCTGC	720
Oy	935	- ACCCTGAGCTGCGACAGAA--CCGATGCCCTTACCCCCAACCCGA--TCCCATCTCTGC	990
Db	721	AACCCCTGAGCTGCGACAGAAACCCGATGCCCTTAAACCCGAACGAAATCCATCTCGGG	780
Oy	991	GAGGCTGCCACAGATGATGCC 1010	
Db	781	GAGGCTGCCACAGATGATGCC 800	

RESULT 14

BX405957

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1126 bp mRNA linear EST 15-MAY-2003

BX405957 Homo sapiens FETAL LIVER Homo sapiens CDNA clone

CSDDM009Y15 5-PRIME, mRNA sequence.

BX405957

BX405957.1 GI:30768575

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1126)

L.I.W.-B., Gruber, C., Jeejee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10000.f For more information about this cluster, see

http://www.genoscope.cns.fr/

cgl-bin/cluster.cgi?seq=CS0AM009CC080P1&cluster=10000.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CS0AM009CC080P1.

Location/Qualifiers

	source	1..1126	/organism="Homo sapiens"
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CSODM009P15"	
		/tissue_type="FETAL LIVER"	
		/dev_stage="fetal"	
		/clone_id="Homo sapiens FETAL LIVER"	
		/note="Torgan: liver; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."	
	BASE COUNT	176 a 360 c 350 g 211 t	29 others
	ORIGIN		
	Query Match	66.8%; Score 717.6; DB 13;	Length 1126;
	Best Local Similarity	98.4%; Pred. No. 2.6e-139;	
	Matches 783; Conservative	4; Mismatches 3;	Indels 6; Gaps 6
QY	1 ATGGAGGACCCCTGTGGCTGCTTCGCGGGATTCCGTGCAGGAAGAAGCGTTTGAG	60	
DB	154 ATGCCAGCACCCCTCTGTGCTGCTCCTCGGGATTCCTGTCAAGAAAGCGATTGAG	213	
QY	61 TTGATGACAATTAGATACCGAGCGTCCGTCCAGAACGAGCTCGAAGTGGCCCCAG	120	
DB	214 TTGATGACAATTAGATACCGAGCGTCCGTCCAGAACGAGCTCGAAGTGGCCCCAG	273	
QY	121 CCCAGACTGCCCCCTGCTGCTGTGGCCCCCTTAGAGCCCACTACTGCTCCAGATCGTCAACT	180	
DB	274 CCCAGACTGCCCCCTGCTGCTGTGGCCCCCTTAGAGCCCACTACTGCTCCAGATCGTCAACT	333	
QY	181 GCTGTGGGCATGCTCCCGTCTTGGGGCCCTATGTCCTCTGAGACCAGAGAGGGCGGG	240	
DB	334 GCTGTGGGCATGCTCCCGTCTTGGGGCCCTATGTCCTCTGAGACCAGAGAGGGCGGG	393	
QY	241 CGGGCTTACCGGGCCCTGCACTGCCCTTACAGGCACTAGTAATCTGCAAGGTGTACCC	300	
DB	394 CGGGCTTACAGGCCCTTGCACATGCCCTTACAGGCACTAGTAATCTGCAAGGTGTACCC	452	
QY	301 GTCCAGGAAGCCCTGCGCCGTGTGAGAGCCCTACGCGGGCGTGGCCCGGCACAAAGCATGTG	360	
DB	453 GTCCAGGAAGCCCTGCGCCGTGTGAGAGCCCTTACGCGGGCGTGGCCCGGCACAAAGCATGTG	512	
QY	361 GCTCGGCCCATGAGAGTCTGTGGTGTATCCAGAGCTCTTACAGCTTTTTCATCTGGAGC	420	
DB	513 GCTCGGCCCATGAGAGTCTGTGGTGTATCCAGAGCTCTTACAGCTTTTTCATCTGGAGC	572	
QY	421 CATGGGGACATGACACAGCTGTGTGGAAAGCCGCCACCGTATCCCTGAGGCTGAGGCTGCC	480	
DB	573 CATGGGGACATGACACAGCTGTGTGGAAAGCCGCCACCGTATCCCTGAGGCTGAGGCTGCC	631	
QY	481 GTGCTCTTCGCGCAGATAGGCAACGCGCCCTGGCGGCACTGTACACAGCAAGGTCTGGTCTGTG	540	
DB	632 GTGCTCTTCGCGCAGATAGGCAACGCGCCCTGGCGGCACTGTACACAGCAAGGTCTGGTCTGTG	691	
QY	541 CGTATATCAAGCTGTGTGCTTTGTCTTTCGCTGACCGTGAAGAGAAAGACTGGTGTCTG	600	
DB	692 CGTATATCAAGCTGTGTGCTTTGTCTTTCGCTGACCGTGAAGAGAAAGACTGGTGTCTG	751	
QY	601 GAGAACTCTGAGAGACTCTGTGTGTCTGACTGTGGGCAATGATTTCCCTGTGGGACAAAGAC	660	
DB	752 GAGAACTCTGAGAGACTCTGTGTGTCTGACTGTGGGCAATGATTTCCCTGTGGGACAAAGAC	811	
QY	661 GCCTGTGCCAGCTTGTGTGGGACCTGAGATACTCAAGCTCAACGGGCTCTATCTCGGGCAAG	720	
DB	812 GCCTGTGCCAGCTTGTGTGGGACCTGAGATACTCAAGCTCAACGGGCTCTATCTCGGGCAAG	870	
QY	721 GCACCCCATGTCTGTGAAGCTGTGGGCTGTCTTACACATGTCTGGCCGACCTTACCCC	780	
DB	871 GSACCCCATGTCTGTGAAGCTGTGGGCTGTCTTACACATGTCTGGCCGACCTTACCCC	927	
QY	781 TTCAGAGACTCGAGC 796		

```

Db      928  TTCAGAGACTCGAGC 943
|||||
RESULT 15
BX363066
LOCUS    BX363066 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION
ACCESSION
BX363066
VERSION  BX363066.1 GI:30380698
KEYWORDS
SOURCE    EST.
ORGANISM  Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 944)
Full-length cDNA libraries and normalization
Unpublished
REFERENCE
AUTHORS  Li, W.B., Gruber, C., Jessup, J. and Polayes, D.
TITLE    Full-length cDNA libraries and normalization
JOURNAL
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10000.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK008B05QPlc1cluster=10000.f. Contact :
Peng Liang Email: filiang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope Sequence ID: CS0DK008B05QPl.
Location/Qualifiers
1..944
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK008YCI0"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT  157 a 315 c 296 g 168 t 8 others
ORIGIN
Query Match      66.2%; Score 711; DB 13; Length 944;
Beebe Local Similarity 99.2%; Pred. No. 5.9e-138;
Matches 744; Conservative 1; Mismatches 2; Indels 3; Gaps 3;

QY      1  ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTTCCCTGTCCAGAGAGAGCGGTTGAG 60
|||||
DB      197  ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTTCCCTGTCCAGAGAGAGCGGTTGAG 256
|||||
QY      61  TTGGATGACAACTTAATACCGAGCGTCCCGTCCAGAAAGAGCTGGAATGGGCCCCAG 120
|||||
DB      257  TTGGATGACAACTTAATACCGAGCGTCCCGTCCAGAAAGAGCTGGAATGGGCCCCAG 316
|||||
QY      121  CCCAGACTGCCCCCTGCTGCTGCTGCTGAGGCCACTTACTCTCCAGATCGTGCAACT 180
|||||
DB      317  CCCAGACTGCCCCCTGCTGCTGCTGAGGCCACTTACTCTCCAGATCGTGCAACT 376
|||||
QY      181  GCTGTGGCCACTGCTCTCCGCTTGGGCGCTATGCTCTCTGTGAGGCCGAGAGAGCGGAG 240
|||||
DB      377  GCTGTGGCCACTGCTCTCCGCTTGGGCGCTATGCTCTCTGTGAGGCCGAGAGAGCGGAG 436
|||||
QY      241  CCGGAGCTACCGGGCGCTGAGCTGCGCTTACAGGCACTGAGATTAAGCTGGAAGGTTACCCC 300
|||||
DB      437  CCGGAGCTACCGGGCGCTGAGCTGCGCTTACAGGCACTGAGATTAAGCTGGAAGGTTACCCC 495
|||||
QY      301  GTCCAGGAAGCCCTGCGCGTGTGAGAGCCCTACGCGCGCTGCCCGGCAAGAGATGTG 360
|||||

```

```

DB      496  GTCCAGGAAGCCCTGCGCGTGTGAGAGCCCTACGCGCGCTGCCCGGCAAGAGATGTG 555
|||||
QY      361  GCTCGGCCCACTGAGGTCCTGCTGCTGTAACCAAGCTCTCTTCACTCGAGCC 420
|||||
DB      556  GCTCGGCCCACTGAGGTCCTGCTGCTGTAACCAAGCTCTCTTCACTCGAGCC 615
|||||
QY      421  CATGGGACATATCACAGCTGTGTGGAGCCGCCCAACCGTATCCCTGAGGCTGAGGCTGCG 480
|||||
DB      616  CATGGGACATATCACAGCTGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
|||||
QY      481  GTGCTCTTCCGCGCAGATGAGCCAGCCGCTGAGGCACTGTACAGAGAGAGAGAGAGAGAG 540
|||||
DB      675  GTGCTCTTCCGCGCAGATGAGCCAGCCGCTGAGGCACTGTACAGAGAGAGAGAGAGAGAG 734
|||||
QY      541  CGTGATCTCAAGCTGTGTGCTTTGTCTTGTGCTGACCGTGAAGAGAGAGAGAGAGAGAG 600
|||||
DB      735  CGTATCTCAAGCTGTGTGCTTTGTCTTGTGCTGACCGTGAAGAGAGAGAGAGAGAGAG 794
|||||
QY      601  GAGAACTT- GAGAGACTCTGTGCTGTGACTGAGGCGCAATGATTCCTGTGTGGACAAAGCA 659
|||||
DB      795  GAGAACTTGGAGAGAGACTCTGTGCTGTGACTGAGGCGCAATGATTCCTGTGTGGACAAAGCA 854
|||||
QY      660  CGCGTGGCCAGGCTTACGTGGAGACTGAGATTAAGCTGACGAGGAGCTTCACTCGAGGCAA 719
|||||
DB      855  CGCGTGGCCAGGCTTACGTGGAGACTGAGATTAAGCTGACGAGGAGCTTCACTCGAGGCAA 914
|||||
QY      720  GGCAGGCGATGTCTGAGAGCTGTGGCGGTGCG 749
|||||
DB      915  GGCAGGCGATGTCTGAGAGCTGTGGCGGTGCG 944
|||||

```

Search completed: January 16, 2004, 16:46:21
 Job time : 2414.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:09:04 (Search time 304.409 Seconds
(without alignment) 9524.031 Million cell updates/sec

Title: US-09-909-474d-1_COPY_49_1122

Perfect score: 1074
Sequence: 1 atgcgagccaccctctcgc.....gagaagctctctatgc 1074

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_GeneSeq_19Jun03.*
1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1070.8	99.7	1077	22	AA506709
2	1069.2	99.6	2059	22	AA159850
3	1069.2	99.6	2092	22	AA158064
4	1069.2	99.6	2116	22	AA130480
5	1051.6	97.9	1083	24	ABN86479
6	956.4	89.1	1085	21	AA261155
7	879.2	81.9	2048	21	AA277866
8	799.2	74.4	972	22	AAH76218

9	597.2	55.6	1076	24	ABL39747	Human NS cDNA sequ
10	597.2	55.6	1076	24	ABL39762	Human NS cDNA sequ
11	426.2	39.7	2562	16	AAQ89817	DNA encoding the A
12	289.4	26.9	534	23	AA586600	DNA encoding novel
13	282.2	26.3	396	22	AA537461	Novel human diagno
14	274.2	25.5	3319	24	ABN86478	Human tribbles hom
15	263.2	24.5	1909	22	AA545035	cDNA encoding nove
16	251.4	23.4	3280	25	ABX63269	Human cDNA #369 di
17	247	23.0	3324	24	AA594913	Human DNA sequence
18	236.2	22.0	690	24	ABT09024	Phase-1 Rat CT gen
19	234	21.8	675	24	ABK83550	Human cDNA differe
20	233.4	21.7	942	25	ABX74428	Human C8FW encodin
21	232.4	21.6	818	25	ACC45125	Human C8FW encodin
22	222.2	20.7	621	25	ACC45125	Human C8FW encodin
23	220.8	20.6	562	21	AACT8131	Human cancer a88oc
24	207.6	19.3	1943	22	AA545223	cDNA encoding nove
25	165	15.4	498	24	ABK65084	Rat sequence diffe
26	163.6	15.2	7282	22	ABA65318	Human immune/haema
27	158	14.7	562	22	ABA61922	Human immune/haema
28	158	14.7	562	22	AAK10231	Human brain expres
29	158	14.7	562	22	AAK36130	Human bone marrow
30	158	14.7	562	22	AAI41844	Probe #10530 used
31	158	14.7	562	23	ABS35823	Human liver single
32	158	14.7	562	24	ABS10267	Human genome-deriv
33	158	14.7	25772	22	ABA81332	Human immune/haema
34	156	14.5	417	22	ABA74417	Human foetal liver
35	156	14.5	417	22	AAK22879	Human brain expres
36	156	14.5	417	22	AAK49052	Human bone marrow
37	156	14.5	417	22	AAK54879	Probe #23565 used
38	156	14.5	417	23	ABS48709	Human liver single
39	156	14.5	417	24	ABS26668	Human genome-deriv
40	142.8	13.3	2038	23	ABL29125	Drosophila melanog
41	108.2	10.1	6988	23	ABL29124	Drosophila melanog
42	85	7.9	1302	17	AA738285	Protein kinase cDN
43	85	7.9	1302	20	AA796447	Human lkb1 coding
44	85	7.9	1302	20	AAK39660	Renal cancer a88oc
45	85	7.9	1302	24	ABA03938	Human STR11 coding

ALIGNMENTS

RESULT 1
AA506709
AA506709 standard; cDNA, 1077 BP.
XX
AC AA506709;
XX
DT 12-SBP-2001 (first entry)
XX
DB Polynucleotide sequence encoding human protein kinase #9.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGB-) SUGB INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX
DR WPI, 2001-343950/36.

PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Claim 1; SEQ ID NO 267; 10078pp; English.
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA162213) with neurotrophic,
 CC immunosuppressant and cytoprotective activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC activation of the activities such as: Immune system suppression,
 CC activating/inhibiting activity, chemoclastic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 SQ Sequence 2092 BP; 405 A; 658 C; 593 G; 436 T; 0 other;
 Query Match 99.6%; Score 1069.2; DB 22; Length 2092;
 Best Local Similarity 99.7%; Pred. No. 1.6e-230;
 Matches 1071; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTTCCTGTCAGAGAAAGCGATTGAG 60
 DB 139 ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTTCCTGTCAGAGAAAGCGATTGAG 198
 OY 61 TTGGATGACAACTTAATACCGAGCGTCCCTGTCAGAAAGAGACTGGAAGTGGCCCCAG 120
 DB 199 TTGGATGACAACTTAATACCGAGCGTCCCTGTCAGAAAGAGACTGGAAGTGGCCCCAG 258
 OY 121 CCCAGACTGCCCCCTGCGCTGTTGCCCTGAGGCCACTCTGCTCAAGATCGTGAAC 180
 DB 259 CCCAGACTGCCCCCTGCGCTGTTGCCCTGAGGCCACTCTGCTCAAGATCGTGAAC 318
 OY 181 GCTGTGGCCACTGCTCCCGTCTTGGGCGCTATGTCCTCTGTAGAGCCCGAGAGGCGGG 240
 DB 319 GCTGTGGCCACTGCTCCCGTCTTGGGCGCTATGTCCTCTGTAGAGCCCGAGAGGCGGG 378
 OY 241 CGGGCTTACCGGGCCCTGTCATGCTTCCCTTACAGGCACTGAGATACCTGCAAGGTGACCC 300
 DB 379 CGGGCTTACCGGGCCCTGTCATGCTTCCCTTACAGGCACTGAGATACCTGCAAGGTGACCC 438
 OY 301 GTTCAGAGAGCCCTGACCGGCTGTAGAGCCCTTACGCGCGGCTGCCCGACAGACATGTG 360
 DB 439 GTTCAGAGAGCCCTGACCGGCTGTAGAGCCCTTACGCGCGGCTGCCCGACAGACATGTG 498
 OY 361 GCTCGGCCCACTGAGGTCTGTAGGTGATCCAGACTCTCTTCAAGCTTTTCTACTCGAAC 420
 DB 499 GCTCGGCCCACTGAGGTCTGTAGGTGATCCAGACTCTCTTCAAGCTTTTCTACTCGAAC 558
 OY 421 CATGGGACATGACAGCTGTGTGCGAAGCGGCAACCTTACCTTGAAGCTTGAAGCTGCC 480
 DB 559 CATGGGACATGACAGCTGTGTGCGAAGCGGCAACCTTACCTTGAAGCTTGAAGCTGCC 618
 OY 481 GTGCTCTTCGCGAGATGAGCACCGGCTGTAGAGCCCTTACGAGACAGACAGGCTGTGCTG 540
 DB 619 GTGCTCTTCGCGAGATGAGCACCGGCTGTAGAGCCCTTACGAGACAGACAGGCTGTGCTG 678
 OY 541 CGTGATCTCAAGCTGTGTGTGCTTTGTCTTGTGCTGACCGGTGAGAGAAAGCTGTGTG 600
 DB 679 CGTGATCTCAAGCTGTGTGTGCTTTGTCTTGTGCTGACCGGTGAGAGAAAGCTGTGTG 738
 OY 601 GAGAACCTGAGAGACTCTGCGGTGTGCTGATGAGGCAAGATGATCCCTGTGAGCAAGCAC 660
 DB 739 GAGAACCTGAGAGACTCTGCGGTGTGCTGATGAGGCAAGATGATCCCTGTGAGCAAGCAC 798
 OY 661 GCGTGGCCAGCTCACTGTGGACTGTGAGATCTGAGCTCAAGGAGCTTATATCTCGGCAAG 720

DB 799 GCGTGGCCAGCTCACTGTGGACTGTGAGATCTGAGCTCAAGGAGCTTATCTCGGCAAG 858
 OY 721 GCGAGGATGTCTGAGAGCTGTGAGCGTGTGCTTTCACCAATGCTGACCGGCACTACCCC 780
 DB 859 GCGAGGATGTCTGAGAGCTGTGAGCGTGTGCTTTCACCAATGCTGACCGGCACTACCCC 918
 OY 781 TTCCAGAGACTCGAGGCTTCTGCTCTTCCGCAATCCCGCGGCGGCTTACCGCTTG 840
 DB 919 TTCCAGAGACTCGAGGCTTCTGCTCTTCCGCAATCCCGCGGCGGCTTACCGCTTG 978
 OY 841 CCGTGAAGCTCTCGGCGGCTGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 DB 979 CCGTGAAGCTCTCGGCGGCTGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
 OY 901 GCTGAACGCTCAAGAGCCAGAGCAAGGATCTCTGACACCTGCTGCTGCAAGACCGGATG 960
 DB 1039 GCTGAACGCTCAAGAGCCAGAGCAAGGATCTCTGACACCTGCTGCTGCAAGACCGGATG 1098
 OY 961 CCTTTAGCCCAACCCGATCCATCTGTGGAGAGCTGCCAGGATGCTCTGATGACTG 1020
 DB 1099 CCTTTAGCCCAACCCGATCCATCTGTGGAGAGCTGCCAGGATGCTCTGATGACTG 1158
 OY 1021 GGGCTGAGCAAGAGCCAG 1074
 DB 1159 GGGCTGAGCAAGAGCCAG 1212
 RESULT 4
 AAF30480
 ID AAF30480 standard; cDNA; 2116 BP.
 XX
 AC AAF30480;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 DB Human protein phosphatase and kinase protein-5 cDNA 1271505CB1.
 XX
 KM Protein phosphatase and kinase protein; PPHK-5; human;
 KM gastrointestinal disorder; immune system disorder;
 KM neurological disorder; cell proliferative disorder; cancer;
 KM diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 161..1237
 FT /tag= a
 FT misc_feature 326..385
 FT /tag= b
 FT /note= "unique fragment"
 XX
 PN W0200120004-A2.
 XX
 PD 22-MAR-2001.
 XX
 PF 14-SEP-2000; 2000MO-US25515.
 XX
 PR 15-SEP-1999; 99US-0154141.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzal Y,
 PI Lu DM;
 XX
 DR WPI; 2001-244811/25.
 DR P-PSDB; AAB20326.
 XX
 PT Novel human protein phosphatase and kinase proteins for diagnosis,
 PT treatment and prevention of gastrointestinal, immune system,
 PT neurological and cell proliferative disorders -
 XX
 PS Claim 5; Page 98; 103pp; English.

also known as stress kinase inhibitor protein (SKIP-1) polypeptide. The htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory signal in a cell. The polypeptide employed in the method is preferably htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3 C, htrb-3 N C, or htrb-3 N C. It is also useful for providing htrb agonist activity for activating an ERK-mediated signal e.g. AP-1-mediated gene activation signal, an estrogen receptor-mediated gene activation signal, an fibroblast growth factor (FGF) induced signal, or a PMA induced signal, in a cell. Htrb modulators are useful for modulating AP-1 mediated inflammatory signal in a cell such as tumor necrosis factor (TNF) induced inflammatory signal, or an interleukin induced inflammatory signal. htrb proteins are useful in screening assays, predictive medicine and in therapeutics or prophylactics. The htrb proteins are useful for screening compounds e.g. for treating and/or preventing diseases caused by abnormal htrb activity, such as rheumatoid arthritis, diabetes, psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction and cancers. The htrb therapeutics are useful for antagonizing interleukin-1 dependent disorders of human placenta, intraventricular hemorrhage, neonatal white matter damage and subsequent cerebral palsy, and inflammation or autoimmune disorders. The present sequence represents a DNA encoding the htrb-3 polypeptide.

Sequence 1083 BP; 176 A; 374 C; 327 G; 206 T; 0 other;

Query Match 97.9%; Score 1051.6; DB 24; Length 1083;

Best Local Similarity 99.1%; Pred. No. 1.3e-226; Mismatches 4; Indels 6; Gaps 1;

Matches 1070; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

```

Oy 1 ATGCAGAGCACCCTCTGCTGCTCTGCGGCTTCTCTCCAGAGAGAGCGTTGAG 60
Db 1 ATGCAGAGCACCCTCTGCTGCTCTGCGGCTTCTCTCCAGAGAGAGCGTTGAG 60
Oy 61 TTGGATGACACTTATGATACGAGCGTCCGCTCCAGAAACGAGCTGAAAGTGGCCCA 120
Db 61 TTGGATGACACTTATGATACGAGCGTCCGCTCCAGAAACGAGCTGAAAGTGGCCCA 120
Oy 121 CCCAGACGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 CCCAGACGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Oy 181 GCTGTGCGCACTGCTCTCCGCTCTTGTGAGCCCTATGCTCTCTGAGAGCCGAGG 240
Db 181 GCTGTGCGCACTGCTCTCCGCTCTTGTGAGCCCTATGCTCTCTGAGAGCCGAGG 240
Oy 241 CGGCGCTTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 CGGCGCTTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Oy 301 GTCAGAGAGCCCTGAGCGCTGAGAGCCCTAGCGGCGCTGCTGAGAGAGATG 360
Db 301 GTCAGAGAGCCCTGAGCGCTGAGAGCCCTAGCGGCGCTGCTGAGAGAGATG 360
Oy 361 GCTCGGCGCACTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 GCTCGGCGCACTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Oy 421 CATGGGAGATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 CATGGGAGATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Oy 481 GTGCTCTTCCGCGCAATGAGCCAGCCGCTGAGCGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTCTTCCGCGCAATGAGCCAGCCGCTGAGCGCTGCTGCTGCTGCTGCTG 540
Oy 541 CGTATCTCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
Db 541 CGTATCTCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
Oy 595 GTGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 654
Db 601 GTGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 660
Oy 655 AAGCAGCGGTGCGCAGCTTACGTGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 714

```

```

Db 661 AAGCAGCGGTGCGCAGCTTACGTGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 720
Oy 715 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
Db 721 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Oy 775 TACCCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834
Db 781 TACCCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Oy 835 GCTTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 894
Db 841 GCTTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Oy 895 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954
Db 901 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Oy 955 CGATGCTGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
Db 961 CGATGCTGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Oy 1015 GACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
Db 1021 GACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080

```

RESULT 6
AAZ61155
ID AAZ61155 standard; CDNA; 1085 BP.

AAZ61155;
30-MAY-2000 (first entry)

CDNA J5503-KS encoding domain VIA to XI of a protein kinase.
Kinase activity; molecular weight marker; isoelectric focusing marker;
peptide fragmentation control; cellular signal transduction; ss.

Homo sapiens.

Key Location/Qualifiers
FT 2..1081
FT /*tag= a
FT /note= "partial sequence"

MO200008180-A2.
17-FEB-2000.

03-AUG-1999; 99WO-US17630.
04-AUG-1998; 98US-0095270.
11-SEP-1998; 98US-0099972.

(IMMUNEX CORP.
Vicea GD, Bird TA, Anderson DM, Marken JS;
WPI; 2000-195584/17.
P-88DB; AAY69157.

New human kinase polypeptides and polynucleotides used as molecular weight markers and as controls for peptide fragmentation -
Claim 1, Page 7; 60pp; English.

The present sequence encodes a partial polypeptide which has kinase activity. The kinase polynucleotides can be used to express the polypeptides, and as probes to identify nucleic acids encoding proteins having kinase activity. The kinase polypeptides and

XX Sequence 2048 BP, 404 A, 619 C, 585 G, 434 T, 6 other;
80 Query Match 81.9%; Score 879.2; DB 21; Length 2048;
Best Local Similarity 92.6%; Pred. No. 7e-188;
Matches 994; Conservative 3; Mismatches 1; Indels 76; Gaps 4;
QY 1 ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTTTCTCTGCGAGAAAGAGCGGTTGGAG 60
DB 156 ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTTTCTCTGCGAGAAAGAGCGGTTGGAG 215
QY 61 TTGGATGACAACTTAAGTACCGAGCGTCCCGTCCAGAAAGAGCTGAAATGGGCCCCAG 120
DB 216 TTGGATGACAACTTAAGTACCGAGCGTCCCGTCCAGAAAGAGCTGAAATGGGCCCCAG 215
QY 121 CCCAGACTGCCCCCTGCTGCTGTTGCCCTGAGGCCCACTAAGCTCCAGATCTGCAACT 180
DB 276 CCCAGACTGCCCCCTGCTGCTGTTGCCCT----- 304
QY 181 GCTGTGGCCACTGCTCCCTCTTGGGCGCTATGCTCTCTGAGGCCCGAGAGGCGCGG 240
DB 305 -----GAGCCCGAGAGGCGCGG 322
QY 241 CGGCGCTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 323 CGGCGCTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380
QY 301 GTCCAGAGAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 381 GTCCAGAGAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 440
QY 361 GCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 441 GCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
QY 421 CATGGGAGCAATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 501 CATGGGAGCAATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 559
QY 481 GTGCTCTTCCGCAATGCGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 560 GTGCTCTTCCGCAATGCGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
QY 541 CGTGAATCTCAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 620 CGTGAATCTCAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679
QY 601 GAGAACCTGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 680 GAGAACCTGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 739
QY 661 GCGTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 740 GCGTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799
QY 721 GCAGCGGATGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 800 GCAGCGGATGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 859
QY 781 TTTCAGAGCTCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 860 TTTCAGAGCTCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 919
QY 841 CCGTGAAGCTCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 920 CCGTGAAGCTCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979
QY 901 GCTGAAGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 980 GCTGAAGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1039
QY 961 CCTTATGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

DB 1040 CCCTTAGTCYCAACCCGATCCCATCTCTGAGAGGCTGCCAGATGCTCCTGATGACTG 1099
QY 1021 GGGCTGAGAGAGCCCGAG 1074
DB 1100 GGGCTGAGAGAGCCCGAG 1153
RESULT 8
AAH76218
ID AAH76218 standard; cDNA, 972 BP.
XX
AC AAH76218;
XX
DT 29-OCT-2001 (first entry)
XX
XX Human kinase PKIN-10 encoding cDNA.
DB
XX
XX PKIN; kinase; cytosolic; immunosuppressive; immunostimulant; human;
KM antileukemic; cardiac; gene therapy; antileukemia therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..972
FT /*tag= a
FT /product= "PKIN-10"
XX
XX MO200160991-A2.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001MO-US05240.
XX
XX 17-FEB-2000; 2000US-0183682.
XX 02-MAR-2000; 2000US-0186559.
XX 09-MAR-2000; 2000US-0186506.
XX 30-MAR-2000; 2000US-0189998.
XX 30-MAR-2000; 2000US-0193851.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Tang YT, Buford N, Gandhi AR, Patterson C, Khan FA, Yue H;
PI Hafalia A, Shih LL, Tribouley CM, Yao MG, Burrill JD, Marcus GA;
PI Zingler KA, Lu DM, Bandman O, Policky JL, Griffin JA, Thornton M;
PI Nguyen DB, Lal P, Walsh RT;
XX
XX MPI; 2001-514771/56.
XX P-PSDB; AAB85791.
XX
XX Isolated human kinase polypeptides useful in the diagnosis, treatment
PT and prevention of cancer; immune disorders and disorders affecting
PT growth and development -
XX
XX Claim 5; Page 125; 126pp; English.
XX
XX The invention provides human kinases (PKIN) and polynucleotides encoding
XX PKIN. The PKIN polypeptides can be expressed using standard recombinant
XX methodology. The PKIN polypeptides, polynucleotides, modulators and
XX specific antibodies are useful in the diagnosis, treatment and prevention
XX of cancer, immune disorders, disorders affecting growth and development,
XX atherosclerosis, and other cardiovascular diseases, and lipid disorders
XX and in the assessment of the effects of exogenous compounds on the
XX expression of nucleic acid sequences of human kinases. The present
XX sequence represents a cDNA encoding a human PKIN-10 polypeptide.
XX
XX Sequence 972 BP, 167 A, 329 C, 294 G, 182 T, 0 other;
QY
Query Match 74.4%; Score 799.2; DB 22; Length 972;
Best Local Similarity 87.6%; Pred. No. 5.5e-170;
Matches 941; Conservative 0; Mismatches 28; Indels 105; Gaps 2;
QY 1 ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTTTCTCTGCGAGAAAGAGCGGTTGGAG 60

```

Db      1  ATGGAGGCAACCCCTTGGCTGCTTCTGCGGGTCCCTGTCCAGGAAGAGCGTTGAG 60
Qy      61  TTGGATGCAACTTATGATACCGAGCGTCCCTGTCAGAAACGAGCTGAAAGTGGCCCA 120
Db      61  TTGGATGCAACTTATGATACCGAGCGTCCCTGTCAGAAACGAGCTGAAAGTGGCCCA 120
Qy      121  CCCAGACTGCCCCCTGCTGTGGCCCTGAGCCCACTGCTGCTGAGATGATGATCACT 180
Db      121  CCCAGACTGCCCCCTGCTGTGGCCCTGAGCCCACTGCTGCTGAGATGATGATGATCACT 180
Qy      181  GCTGTGACCACTGCTCCCTGCTTGGAGCCCTATGCTCTCTGAGAGCCCGAGAGGCG 240
Db      181  GCTGTGACCACTGCTCCCTGCTTGGAGCCCTATGCTCTCTGAGAGCCCGAGAGGCG 240
Qy      241  CGGCGCTTACCGGGCCCTGCTGACCTGCTTACAGAGCACTGATATCTGCAAGGTAT 300
Db      241  CGGCGCTTACAGGCGCTTCACTGCTTACAGAGCACTGATATCTGCAAGGTATACCC 300
Qy      301  GTCCAGAAAGCCCTGCGCTGAGAGCCCTACGCGGCGCTGCCCGGCAAGAGATG 360
Db      301  GTCCAGAAAGCCCTGCGCTGAGAGCCCTATGCGGCGCTGCCCGGCAAGAGATG 360
Qy      361  GCTCGGCCCACTGAGGCTGCTGCTGATCCCACTTCTCTACGCTTTTCACTTGAC 420
Db      361  GCTTGGCCCACTGAGGCTGCTGCTGATCCCACTTCTCTACGCTTTTCACTTGAC 420
Qy      421  CATGGGAGCATGCAAGCTGTGTGGAGAGCGGCAACCGATCCCTGAGCTGAGGCTGC 480
Db      421  CATGGGAGCATGCAAGCTGTGTGGAGAGCGGCAACCGATCCCTGAGCTGAGGCTGC 480
Qy      481  GTGCTCTTCCGCAAGATGAGCCGCGCTGCGCACTGCAACAGCAAGCTGTGCTG 540
Db      481  GTGCTCTTCCGCAAGATGAGCCGCGCTGCGCACTGCAACAGCAAGCTGTGCTG 540
Qy      447  -----GCACAGCCCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
Db      447  -----GCACAGCCCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
Qy      541  CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db      541  CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy      475  -----ACACGGAAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
Db      475  -----ACACGGAAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
Qy      601  GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db      601  GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy      496  GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
Db      496  GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
Qy      661  GCGTCCAGGCTTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db      661  GCGTCCAGGCTTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy      556  GCGTCCAGGCTTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
Db      556  GCGTCCAGGCTTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
Qy      721  GCAAGCGATGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      721  GCAAGCGATGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy      616  GCAGCAATGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Db      616  GCAGCAATGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Qy      781  TTCCAGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db      781  TTCCAGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy      676  TTCCAGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
Db      676  TTCCAGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
Qy      841  CCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db      841  CCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy      736  CCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
Db      736  CCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
Qy      901  GCTGAAAGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db      901  GCTGAAAGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy      796  GCTGAAAGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 855
Db      796  GCTGAAAGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 855
Qy      961  CCCTTAAAGCCCAACCGATCCATCTCTGAGAGGCTGCTGCTGCTGCTGCTGCT 1020
Db      961  CCCTTAAAGCCCAACCGATCCATCTCTGAGAGGCTGCTGCTGCTGCTGCTGCT 1020
Qy      856  CCCTTAAAGCCCAACCGATCCATCTCTGAGAGGCTGCTGCTGCTGCTGCTGCT 915
Db      856  CCCTTAAAGCCCAACCGATCCATCTCTGAGAGGCTGCTGCTGCTGCTGCTGCT 915
Qy      1021  GGGCTGAGCAAGCCAGGAAAGAGAGGAGAGCAAGAGAGTGTCTGTATGCG 1074
Db      1021  GGGCTGAGCAAGCCAGGAAAGAGAGGAGAGCAAGAGAGTGTCTGTATGCG 1074
Qy      916  GGGCTGAGCAAGCCAGGAAAGAGAGGAGAGCAAGAGAGTGTCTGTATGCG 969
Db      916  GGGCTGAGCAAGCCAGGAAAGAGAGGAGAGCAAGAGAGTGTCTGTATGCG 969

```

```

ID      1  ABLJ9747 standard; cDNA; 1076 BP.
XX
AC      ABLJ9747;
XX
DT      10-MAY-2002 (first entry)
XX
DE      Human NS cDNA sequence SEQ ID NO:57.
XX
KW      Human; cytosolic; osteopathic; gynaecological; neuroprotective;
KW      antirheumatic; antirheumatic; antipsoriatic; ophthalmological; anti-HIV;
KW      vasorelaxant; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
KW      anorectic; muscular; antidiabetic; antidiabetic; antidiabetic;
KW      antidiabetic; hypotension; antidiabetic; immunomodulator; cardiac;
KW      antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
KW      gastrointestinal; virucide; antitumor; cerebroprotective; neurologic;
KW      contraceptive; vaccine; gene therapy; cancer; osteoporosis; dyatonia;
KW      endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW      rheumatoid arthritis; cataract; retinosis; atherosclerosis; glaucoma;
KW      inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW      infertility; cardiovascular disease; coagulation disease; hypertension;
KW      leishmania; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW      diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW      gastric ulcer; Alzheimer's disease; gene; ss.
XX
OS      Homo sapiens.
XX
XX      WO200206315-A2.
XX
XX      24-JAN-2002.
XX
XX      17-JUL-2001; 2001WO-IL00653.
XX
XX      18-JUL-2000; 2000IL-0137345.
XX
XX      15-DEC-2000; 2000IL-0140354.
XX
XX      (COMP-) COMPUGEN LTD.
XX
XX      Mintz L, Freilich S, Bernstein J;
XX      WPI, 2002-155037/20.
XX      DR      P-PsDB; ABB06093.
XX
XX      One hundred and twenty eight novel nucleic acid sequences, useful for
XX      treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
XX      Claim 1; Page 107; 290pp; English.
XX
XX      ABLJ9691 to ABLJ9818 represent novel human nucleic acid sequences
XX      encoding the proteins given in ABB06037 to ABB06164. The novel sequences
XX      (NS) can have cytosolic, osteopathic, gynaecological, neuroprotective,
XX      antirheumatic, antirheumatic, antipsoriatic, ophthalmological, virucide,
XX      vasorelaxant, antidiabetic, antidiabetic, antidiabetic, antidiabetic,
XX      anorectic, muscular, antidiabetic, antidiabetic, antidiabetic,
XX      antidiabetic, hypotension, antidiabetic, immunomodulator, cardiac,
XX      immunomodulator, antidiabetic, antidiabetic, antidiabetic, antidiabetic,
XX      antidiabetic, gastrointestinal, neurologic, cerebroprotective,
XX      neurotropic and contraceptive activities. The NS can be used in vaccines and
XX      gene therapy and antineoplastic therapy. Nucleic acids, expression vectors and
XX      antibodies from the present invention can be used for treating and
XX      diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
XX      diseases, dyatonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
XX      cataracts, retinosis, atherosclerosis, inflammation, skin disorders,
XX      glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
XX      disease, coagulation disease, leishmania, hypertension, asthma, immune
XX      disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
XX      depression, schizophrenia, viral disease, gastric ulcers, stroke,
XX      Alzheimer's disease and as a contraceptive.
XX
XX      Sequence 1076 BP; 172 A; 352 C; 346 G; 194 T; 12 other;

```



```

Oy 361 GCTGGCCCACTGAGAGTCTGCTGTAACCCAGCTCTCTAGACGCTTTTCACTGGACC 420
Db 689 GCTGGCCCACTGAGAGTCTGCTGTAACCCAGCTCTCTAGACGCTTTTCACTGGACC 748
Oy 421 CATGGGACATGACAGAGCTGTGGAGAGCCGCAACCGATCTCTAGAGCTGAGCTGCC 480
Db 749 CATGGGACATGACAGAGCTGTGGAGAGCCGCAACCGATCTCTAGAGCTGAGCTGCC 808
Oy 481 GTGCTCTTCCGCAATGAGCCACCCGCTGGGCACTGTACACAGACGATCTGCTCTG 540
Db 809 GTGCTCTTCCGCAATGAGCCACCCGCTGGGCACTGTACACAGACGATCTGCTCTG 868
Oy 541 GGTGATCTCAAGCTGTGTGCTTTGTCTTGTGCTGACCGTAGAGGAGAGAGCTGAGCTG 600
Db 869 GGTGATCTCAAGCTGTGTGCTTTGTCTTGTGCTGACCGTAGAGGAGAGAGCTGAGCTG 927
Oy 601 GAGAACCTGAGAGACTCTGCTGCTGTGACTGAGCTGGGC 634
Db 928 GAGA--CCTGAGAGACTGCTGCTGCTGAGCTGGGC 959

```

RESULT 11

AA089817/c
ID AA089817 standard, DNA; 2562 BP.

AA089817:
AC 25-MAR-2003 (updated)
DT 27-NOV-1995 (first entry)
XX

DE DNA encoding the AUP1 polypeptide.

KM AU-rich element RNA-binding factor-1; proto-oncogene; c-myc; c-myb;
KW c-fos; cancer; ss.

OS Homo sapiens.

PH Key Location/Qualifiers
FT CDS 246..1109
FT /*tag= a

XX MO9511695-A1.

XX PD 04-MAY-1995.

XX PP 25-OCT-1994; 94WO-US12273.

XX PR 29-OCT-1993; 93US-0146421.

XX PA (UYMA-) UNIV WAKE FOREST.

XX PI Brewer G;

XX WP1; 1995-178647/23.

XX DR P-PSDB; AAR74201.

XX PT AUP1 polypeptide and DNA encoding it - limits the expression of a
XX proto-oncogene that, when over expressed, leads to cancer.

XX PS Claim 3, Fig 8, 87p; English.

XX The sequence is that of the DNA encoding AUP1 (AU-rich element RNA-
XX binding factor-1). AUP1 can be used to treat humans with low levels
XX of the AUP1 gene, so limiting the expression of a proto-oncogene (
XX pref. c-myc, c-myb or c-fos) which when over expressed, leads to
XX cancer.
XX See also AA089818-20.
XX (Updated on 25-MAR-2003 to correct PN field.)

XX SO Sequence 2562 BP; 663 A; 608 C; 788 G; 503 T; 0 other;

Query Match 39.7%; Score 426.2; DB 16; Length 2562;
Best Local Similarity 93.8%; Pred. No. 3.4e-86;

```

Matches 487; Conservative 0; Mismatches 28; Indels 4; Gaps 4;
Oy 556 TGTGCTTTGTCTTCGCTGACCGCTGAGAGAGAAAGCTGTGCTGAGAACTGAGAGC 615
Db 1637 TGACCCCTTCTGTCTTCCCAATGCTCCAGAAAGAGCTGTGCTGAGAACTGAGAGC 1578
Oy 616 TCCGCGCTGCTGACCTGGGCAATGATTCCTGTGTGGACAAACAGCCGCTGCCACTAC 675
Db 1577 TCTGCGGTGCTGACCTGGGCAATGATTCCTGTGTGGACAAACAGCCGCTGCCACTAC 1519
Oy 676 GTGGACCTGAGATATCAAGCTCAAGGACCTCAATATCGAGGAGAGAGCCGATGTCTGA 735
Db 1518 GTGGACCTGAGATATCAAGCTCAAGGACCTCAATATCGAGGAGAGAGCCGATGTCTGA 1459
Oy 736 AGCTTGAGCGTGGCGCTCTTCAACATGCTGAGCGGCACTACCCCTTCCAGAGCTCGAG 795
Db 1458 AGCTTGAGCGTGGCGCTCTTCAACATGCTGAGCGGCACTACCCCTTCCAGAGCTCGAG 1399
Oy 796 CCTGTCTGTCTTTCGGGAATTCGCGCGGAGGCTTAAGCCTTGTCCAGAGCCTCTTG 855
Db 1398 CCTGTCTGTCTTTCGGGAATTCGCGCGGAGGCTTAAGCCTTGTCCAGAGCCTCTTG 1339
Oy 856 GCGCCCTGCGCGCTGTCTGATTGCTGCTGCTCTTGTGCGGAGCCAGCTGAAGGCTCA 915
Db 1338 CCGTGGCGCGCTGTCTGATTGCTGCTGCTCTTGTGCGGAGCCAGCTGAAGGCTCA 1279
Oy 916 GCCACAGGCAATCTCTGACACCCCTGAGCTGCGAGACAGACCCGATGCCCTTACGCCAAC 975
Db 1278 G-CACAGGCAATCTCTGACACCCCTGAGCTGCGAGCA-GACCGGATGCCCTTACGCCAAC 1221
Oy 976 CGATCCCATCTCTGGAGAGCTGCCAGAGTGTCTCTGATGGAATGAGGCTGAGACGAGC 1035
Db 1220 CGATCCCATCTCTGGAGAGCTGCCAGAGTGTCTCTGATGGAATGAGGCTGAGACGAGC 1162
Oy 1036 AGGGAAGAGAGAGAGACAGAGAAAGTGTCTGTATGAGC 1074
Db 1161 AGGGAAGAGAGAGAGACAGAGAAAGTGTCTGTATGAGC 1123

```

RESULT 12

AA568600
ID AA568600 standard, cDNA; 534 BP.

XX AA568600;

XX DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #4404.

KM Human, chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PP 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSB-) HYSBQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WP1; 2001-639362/73.

XX DR P-PSDB; ABG04413.

XX New isolated polynucleotide and encoded polypeptide; useful in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess

FT	CDS	282..1400
PT		/tag= a
PT		/product= "htrb-1"
PT		/note= "ribbles polypeptide"
XX		
PX		
PN		WO200253743-A2.
PD		
PP		11-JUL-2002.
XX		
PR		08-JAN-2002; 2002WO-US00070.
PA		08-JAN-2001; 2001US-260294P.
PI		(INTE-) INTERLEUKIN GENETICS INC.
XX		
DW		Dower S, Quamstrom E, Kiss-Toch E;
DR		WPI; 2002-590635/63.
XX		P-Psdb; ABB80975.
PT		
PT		Novel isolated human tribbles homologue-1 polypeptide for inhibiting
AP		AP-1-mediated inflammatory signal in a cell, and activating
ERK		ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
XX		
PS		Claim 5, Fig 10A; 131pp; English.
CC		The invention provides an isolated human tribbles homologue-1 (htrb-1,
CC		also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC		htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC		signal in a cell. The polypeptide employed in the method is preferably
CC		htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or
CC		htrb-3 N C. It is also useful for providing htrb agonist activity for
CC		activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC		signal, an estrogen receptor-mediated gene activation signal, an
CC		fibroblast growth factor (FGF) induced signal, or a p38 induced signal,
CC		in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC		inflammatory signal in a cell such as tumor necrosis factor (TNF)
CC		induced inflammatory signal, or an interleukin induced inflammatory
CC		signal. htrb proteins are useful in screening assays, predictive medicine
CC		and in therapeutics or prophylactics. The htrb proteins are useful for
CC		screening compounds e.g. for treating and/or preventing diseases caused
CC		by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
CC		psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
CC		and cancers. The htrb therapeutics are useful for antagonizing
CC		interleukin-1 dependent disorders of human placenta, intraventricular
CC		hemorrhage, neonatal white matter damage and subsequent cerebral palsy,
CC		and inflammation or autoimmune disorders. The present sequence represents
CC		a cDNA encoding the htrb-1 polypeptide.
XX		
SQ		Sequence 3319 BP; 782 A; 909 C; 800 G; 827 T; 1 other;
Query Match	25.5%; Score 274.2; DB 24; Length 3319;	
Beat Local Similarity	60.3%; Pred. No. 4, 6e-52;	
Matches 453; Conservative	0; Mismatches 298; Indels 0; Gaps 0	
OY	194 CCTCCCGCTTGGGCGCCTATGTCCCTTCGTGAACCCGAGAGGGCGGCGCTACTACGGG	253
DB		
544	CGAGCGCATGCCGACCTAACCTGCCTGCGCCCTTAGCCAGCGCAGCACTGTGTCCCGG	603
OY	254 CCCTGCATCGCCTTAACAGCACATAAGTATTACCTGTGAAGGTATACCCGTCAGAAAGCC	313
DB	604 CGTGTCATCCACACTGGAACGCGAGTCGCGCTGCAAGGTGTTTCCATTAAACACTACC	663
OY	314 TGCGCTGCTGGAAGCCCTTAAGCGGCGGTGCCCCGCGACAGAAGATGTGGTGGGCCACTG	373
DB	664 AGAACAATAATCAGGCTTACATTCACAGCTGCACATTCGACAGCAATTACTGGCATTTGGG	723
OY	374 AGGTCTGGCTGTATCCAGAGCTCTCTACGCTTTTTCACACTGTGAAGCCATGGGAGCATGC	433
DB	724 AAGTGATCTTTGGGGAAACCAAGGCTATGTCTTTTGGAAAGACTTTGGGAGCATGC	783
OY	434 ACAGCTGTGTCGAAGCCGCCACGTAATCCTTGAGCTGTGAGCTGCGGTCTTTTCGCGC	493

Db	784	ACTCTTATGTGTGCGAAGCCGGAGAGAGCGTCCGGAGAGAGAGAGCCGCCCGCTCTTTCAAGC	843
Qy	494	AGATGATGCCACCGCCCTGGCGCATGTCAACAGACAGTGTGTCTCTGCTGATCTCAAGC	553
Db	844	AGATTGTCTCCGCCGTGGCCCATCTGACACAGTCAAGCATGTGTGTGGGGACCTGAAGC	903
Qy	554	TGTGTGCTTTGTCTTGTGGCTGACCGGTGAGAGAGAGAAAGCTGTGTGTGAGAACTGTGAG	613
Db	904	TTAGAGAGTTGTCTTCTTCCACGAGAGAGAGAACCCAGCTTAGACTAGAAAGTTAGAAAG	963
Qy	614	ACTCTGTGTGTCTGACTGTGGGCGAGATGATTTCCCTGTGTGAGACAGACCGCTGCCAGCT	673
Db	964	ACACACACATTAATGAAGGGAGAGATGATGTCTTTGTGAGACAAACATGTGCTGCCAGCT	1022
Qy	674	ACGTGGGACCTGAGATACTCAGCTCACGGGCGCTCATTACTCGGGCAAGGACCGCATGTCT	733
Db	1024	ACGTGAGCCCTCGAATCTCTCAACACACACTGAGGACCTACTCCGAAAGGCTGCGACGTTT	1082
Qy	734	GGAAGCTGGGGGTGGCGCTCTTTCACCAATGTGTGGCGGACCACTAACCCCTTCCAGGACTGG	793
Db	1084	GGAGCTGTGGGGGTATGTCTTACACCTCTTGTGTGAGACATATCCCTTCCATAGACTGAG	1144
Qy	794	AGCCTGTCTGTCTTTCGGCAAGATCCGCGCGGGGCTTAGCCTTGTGCTGACAGGCTCT	853
Db	1144	ACCCAGTGGCCCTTTTCTCCAAATTTGGGCGGTGACAGTTTGTGATTTCTTGAGACATATT	1202
Qy	854	CGGCGCCCTGGCCCGGTGTCTGTGTTTGGCTGCTCTTGTCTGGAGACGACTGAAGGCTCA	913
Db	1204	CCCCCAAGCCAGGTGCTCTCATTTGCGAGCGCTTTGAGACGGGAGCCCTCCGAGAGACTCA	1262
Qy	914	CAGCCACAGGCAATCTCTCTGACACCCCTGAGT 944	
Db	1264	CTGCCCCGAGATCTTACTGCAACCCCTGATT 1294	

RESULT 15
 AAS45035
 ID AAS45035 standard; cDNA; 1909 BP.
 XX AAS45035;
 AC
 DT 18-DEC-2001 (first entry)
 DB
 XX cDNA encoding novel human secretory protein, Seq ID No 116.
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischemic-reperfusion injury; hematopoietic; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen; se.
 XX Homo sapiens.
 OS
 XX
 PN MO200166689-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001MO-US04942.
 XX
 PR 07-MAR-2000; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 XX
 XX
 XX (HYSB-) HYSBO INC.
 XX

PI Tang YT, Liu C, Abundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P,
 PI Zhao Q, Yang Y, Dmianac RT, Zhang J, Chen R, Xue AJ, Wang J,
 DR WPI, 2001-589934/66.
 XX P-PsDB; AAU28135.

PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 prepared from various human tissues, for diagnosis and treatment of
 cancer, neurological, inflammatory, and autoimmune disorders -
 PT
 XX
 PS Claim 1; SEQ ID No 116; 107pp; English.

CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythm or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamin, mineral, provides
 CC anabolic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAS4920-AAS45295 represent novel human secreted protein
 CC coding sequences of the invention.
 CC
 XX

8Q Sequence 1909 BP; 472 A; 535 C; 435 G; 467 T; 0 other;

Query Match 24.5%; Score 263.2; DB 22; Length 1909;

Best Local Similarity 60.7%; Pred. No. 1.2e-49; Mismatches 288; Indels 1; Gaps 1;

Matches 447; Conservative 0; Mismatches 288; Indels 1; Gaps 1;
 QY 210 CTATGCTCTCTGAGCCCGAGAGAGGCGGCGCTACCGGCGCTTGACCTGCTTAC 269
 DB 103 CTACCTGCTGCTGCGCCCTTAGCGGAGCGAGCATGTGTCCGGGCGTGTGATCCACAC 162
 QY 270 AGGCACTGATTAAC-CTGCAAGGTGTACCCGCTCAGAGAACCTTGCGCTGAGAC 328
 DB 163 TGGAGCGCGAGTGTGCTGCAAGGTGTTCCATTAAACATPACAGAGCAAAATCAAGGC 222
 QY 329 CCTACGCGGCGTGCCTCCCGCAAGCATGTGACTCGGCGCACTGAGGCTTGCTGCTGTA 388
 DB 223 CTACATCTCAGCTGCTGCTGCAACGACATTAAGTGTGAGATGATCTTGGGG 282
 QY 389 CCCAGCTCTCTACGCTTTTCTCACTCGGACCCATGGGAGCATGACAGCTGTGGAA 448
 DB 283 AAACCAAGGCTATGTCTTCTTGGAGAGGACTTTGGGAGCATGCACTCTATGAGCA 342
 QY 449 GCGGCAACCTTATCTCTGAGCTGAGGCTGCGGCTCTTCCGCAAGTGGCAACGCGCC 508
 DB 343 GCGGAAAGGCTGCGGAGAGAGAGCGCGCGCTCTTCAAGGAGATTGTCTCGCGCG 402
 QY 509 TGGCGACTGTGACGAGCAGGCTGTGCTCTGCTGATCTCAAGCTGTGTGCTTGTCT 568

DB 403 TGGCCACCTGCGCACCAAGTCAGCATGTGTGCGGAGCCTGAAAGCTTAGAAATTCGTCT 462
 QY 569 TGGCTGACCGTGAAGAGAGAAAGCTGTGTGTGAGAACTGTGAGACCTTCCTGTGTGA 628
 DB 463 TCTCCAGGAGAGAGAAAGCCAGCTTAGCTAGAAAGCTTAGAAGACACACATTAATGA 522
 QY 629 CTGGGCGAGATGATTCCTGTGTGAGCAAGACGCTGTCCAGCTTACGTGAGACCTGAG 688
 DB 523 AGGGGAGAAAGATGATCTTGTGTGACAAACATGAGGCTGCGCCACCTTACGTGAGCCCTGAG 582
 QY 689 TACTGAGTCAAGGCGCTCATATCTGCGGCAAGGCGAGATGTGTGAGGCTGTGGCGTGG 748
 DB 583 TCTTAACACCACTGGAACCTAATCTCGGAAGGCTGCGGACCTTGGAGCTGTGGGAGTA 642
 QY 749 CGCTCTTACACATGTGTGCGCGGCGCACTAACCTTCCAGAGACTGAGAGCTGTCTGTCT 808
 DB 643 TGTCTTACACCTTGTGTGTGAGACATACCTTCCATGACTCAAGACCCAGTGCCTTT 702
 QY 809 TCGGAGAGATCGCGCGCGGCGCTTACGCTGTGCAAGGCTTGTGCGCTTGTGCGCT 868
 DB 703 TCTCAAAATTCGCGGTGAGACAGTTCTGCAATTCGTAGACATTTCCCAAGCCAGGT 762
 QY 869 GTGTGCTTGTGCTGCTCTTGTGTGAGGCGGAGCTGAAGGCTCAAGGACAGGATCC 928
 DB 763 GCTCATTTGCGAGCTCTTGTGAGACGAGGCGCTTCCGAGAGACTCACTGCGCCGAGATCC 822
 QY 929 TCTGCAACCTGTGCT 944
 DB 823 TACTGCAACCTGTGCT 838

Search completed: January 16, 2004, 11:33:02
 Job time : 311.409 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:11:12, Search time 4024.84 Seconds

(without alignments) 10916.435 Million cell updates/sec

Title: US-09-909-474d-1_COPY_49_1122

Perfect score: 1074

Sequence: 1 atggagagcaccacccctctgc.....gagagtgctctctgtagc 1074

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenEmbl:

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pac:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_ry:*
- 12: gb_un:*
- 13: gb_vl:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pac:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_ry:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hcg_hum:*
- 31: em_hcg_inv:*
- 32: em_hcg_other:*
- 33: em_hcg_mus:*
- 34: em_hcg_dyn:*
- 35: em_hcg_rod:*
- 36: em_hcg_mam:*
- 37: em_hcg_vrt:*
- 38: em_ay:*
- 39: em_hcg_hum:*
- 40: em_hcg_mus:*
- 41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1070.8	99.7	1077	6	AX166518
2	1070.8	99.7	2121	9	BC019363
3	1070.8	99.7	2283	9	BC027484
4	1070.8	99.7	2488	9	AY247738
5	1069.2	99.6	2116	6	AX099934
6	1066	99.3	2257	6	AK026945
7	1051.6	97.9	1083	6	AK572896
8	1051.6	97.9	1083	6	AF250311
9	799.2	74.4	972	6	AX224734
10	656.6	61.1	1256	10	AF358868
11	611.2	56.9	2004	10	AB020967
12	597.2	55.6	1076	6	AX364906
13	597.2	55.6	1076	6	AX364921
14	512.4	47.7	1969	10	BC012955
15	495	46.1	153170	9	HS110307
16	426.2	39.7	2559	6	AR274998
17	426.2	39.7	2559	9	HS002019
18	426.2	39.7	2562	6	I24432
19	297	27.7	75131	2	AC023022
20	296.6	27.6	191503	2	AL928568
21	286.8	26.7	235947	2	AC133773
22	286.8	26.7	282849	2	AC132555
23	282.2	26.3	396	6	AX245589
24	274.2	25.5	3319	6	AX572894
25	274.2	25.5	3319	9	AF250310
26	272.6	25.4	3317	9	AF205437
27	272.6	25.4	3338	9	HSW803696
28	261.4	24.3	2514	10	BC006800
29	258.2	24.0	1244	10	AF358866
30	254	23.6	1682	4	CRC5FW
31	253	23.6	2156	9	BC002637
32	253	23.6	3062	9	AY245544
33	253	23.6	4221	9	D87119
34	250.6	23.3	1102	10	AF205438
35	247	23.0	3324	6	AX281759
36	244.4	22.8	1032	4	CRC5FW
37	241.2	22.5	1496	10	AF358867
38	241.2	22.5	3244	10	BC034338
39	236.2	22.0	690	6	AX525590
40	234	21.8	675	9	HSC8FW
41	222.2	20.7	2667	10	BC027159
42	217.8	20.3	1747	5	AY247742
43	210.8	19.6	2136	5	AY254200
44	165	15.4	498	6	AX401315
45	163.6	15.2	123529	9	AF179296

ALIGNMENTS

RESULT 1	AX166518	1077 bp	DNA	linear	PAT 22-JUN-2001
LOCUS	AX166518				
DEFINITION	Sequence 9 from Patent WO0138503.				
ACCESSION	AX166518				
VERSION	AX166518.1				
KEYWORDS	GI:14546863				
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	1 Plozman,G.D., Whyte,D., Manning,G.S., Sudareanam,S.S., Martinez,R., Planagan,P. and Clary,D.S.				
TITLE	Novel human protein kinases and protein kinase-like enzymes				

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>
 Series: IRAX Plate: 52 Row: F Column: 11.
 Location/Qualifiers
 1. 2283

CDS

BASE COUNT ORIGIN	463 a	702 c	656 g	462 t
----------------------	-------	-------	-------	-------

Query Match	99.74;	Score 1070.8;	DB 9;	Length 2283;
Best Local Similarity	-99.84;	Pred. No. 5.6e-180;		
Matches 1072;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	1	ATGCGAGCACCCCTGCTGGCTGCTCTGGGGTTCCCTGTCAGAGAAAGACGGTTGGAG	60
Db	295	ATGCGAGCACCCCTGCTGGCTGCTCTGGGGTTCCCTGTCAGAGAAAGACGGTTGGAG	354
Qy	61	TTGGATGACAACTTAGATACCGAGCGTCCCTCCAGAAAAGAGCTCCAGATGGGCCCGAG	120
Db	355	TTGGATGACAACTTAGATACCGAGCGTCCCTCCAGAAAAGAGCTCCAGATGGGCCCGAG	414
Qy	121	CCGAGACTGACCCCTGCTGTGGCCCTGAGGCCACTAGCTGCTCCAGATCGTGAACCT	180
Db	415	CCGAGACTGACCCCTGCTGTGGCCCTGAGGCCACTAGCTGCTCCAGATCGTGAACCT	474
Qy	181	GCTGTGGCACTGCTCCCGTCTTGGGCCCTATGTCTCTTGAGACCAGAGAGGGCGGG	240
Db	475	GCTGTGGCACTGCTCCCGTCTTGGGCCCTATGTCTCTTGAGACCAGAGAGGGCGGG	534
Qy	241	CGGGCCCTACCGGGGCCCTGCACTGGCCCTTACAGGCACTAGATATACCTGGAAGTGAACCC	300
Db	535	CGGGCCCTACCGGGGCCCTGCACTGGCCCTTACAGGCACTAGATATACCTGGAAGTGAACCC	594
Qy	301	GTCAGAGAAAGCCCTGACCCTGAGCCCTACACGCGAGGCTGACCCCGCAAGAGCATGTG	360
Db	595	GTCAGAGAAAGCCCTGACCCTGAGCCCTACACGCGAGGCTGACCCCGCAAGAGCATGTG	654
Qy	361	GCTCGGCCCACTGAGGTCTGGCTGTGATCCAGACTCTTACGCGCTTTTCACTCGAAC	420
Db	655	GCTCGGCCCACTGAGGTCTGGCTGTGATCCAGACTCTTACGCGCTTTTCACTCGAAC	714
Qy	421	CATGGGAGCATGACAGCCTGCTGGTGGAGAGCGCACCGTATCCCTGAGCCCTGAGGCTGGC	480
Db	715	CATGGGAGCATGACAGCCTGCTGGTGGAGAGCGCACCGTATCCCTGAGCCCTGAGGCTGGC	774
Qy	481	GTCGCTTTCGCGCAGATGGCCACCGCCCTGGCGCACTGTCAACAGCAAGTCTGTGCTGTG	540
Db	775	GTCGCTTTCGCGCAGATGGCCACCGCCCTGGCGCACTGTCAACAGCAAGTCTGTGCTGTG	834
Qy	541	CGTGAATCAAGCTGTGTGCTTTGTGTGTGCTGACGCTGAGAGAGAAAGCTGTGCTGTG	600
Db		CGTGAATCAAGCTGTGTGCTTTGTGTGTGCTGACGCTGAGAGAGAAAGCTGTGCTGTG	

Db	835	CGGATCTCAAGCTGTGTGCTTTGTCTTGCTGACCGGTGAGAGGAAAGACTGGTGTCTA	894
Qy	601	GAGAACTTGGAGAGCATCTCTGGGCGTGGACTGGGCGCAGATGATATCCCTGGTGGGACAAGCAC	660
Db	895	GAGAACTTGGAGAGCATCTCTGGGCGTGGACTGGGCGCAGATATATCCCTGGTGGGACAAGCAC	954
Qy	661	GCGTGCCCAAGCCTTACGTGGAGCCTGAGATATCATAGCTCAAGGAGCTTATATCTGGGCAAG	720
Db	955	GCGTGCCCAAGCCTTACGTGGAGCCTGAGATATCATAGCTCAAGGAGCTTATATCTGGGCAAG	1014
Qy	721	GCAAGCCGATGTCTGGAGAGCTTGGGCGGTGGGCGCTTTACCAATGTCTGGCGGGCCACTAACCC	780
Db	1015	GCAAGCCGATGTCTGGAGAGCTTGGGCGGTGGGCGCTTTACCAATGTCTGGCGGGCCACTAACCC	1074
Qy	781	TTCCAGGACTCGGAGCGCTGTCTGTCTTCTTGGGCAAGATCCGCGCGGAGGAGCTACGCTTGG	840
Db	1075	TTCCAGGACTCGGAGCGCTGTCTGTCTTCTTGGGCAAGATCCGCGCGGAGGAGCTACGCTTGG	1134
Qy	841	CTTGCAAGCCTCTTGCGGCCCTCTGGCGCGTGTCTGGATTGCTGCTCTTCTGTGGAGGCCA	900
Db	1135	CTTGCAAGCCTCTTGCGGCCCTCTGGCGCGTGTCTGGATTGCTGCTCTTCTGTGGAGGCCA	1194
Qy	901	GCTGGAACGGCTCACAGCCACAGGAGATCTCTGTGCAACCCCTGGGCTGGACAGAGACCCGATG	960
Db	1195	GCTGGAACGGCTCACAGCCACAGGAGATCTCTGTGCAACCCCTGGGCTGGACAGAGACCCGATG	1254
Qy	961	CCCTTAGCCCCAACCCGATCCATCTCTGGAGAGGCTGCCAGGTTGATCCCTGATGGACGTG	1020
Db	1255	CCCTTAGCTCAAAACCCGATCCATCTCTGGAGAGGCTGCCAGGTTGATCCCTGATGGACGTG	1314
Qy	1021	GGGCTGGACGAAGCCAGGGAGAGGAGGGGAGACAGAGAACTGGTTCTGTATGGAC	1074
Db	1315	GGGCTGGACGAAGCCAGGGAGAGGAGGGGAGACAGAGAACTGGTTCTGTATGGAC	1368

RESULT	4
LOCUS	AY247738
DEFINITION	Homo sapiens TRB3 protein mRNA, complete cds.
ACCESSION	AY247738
VERSION	AY247738.1
KEYWORDS	GI:30025661
SOURCE	.
ORGANISM	Homo sapiens (human)
	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Suetheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 2488) Shan, Y.X., and Yu, L. Direct Submission Submitted (03-MAR-2003) School of Life Science, Institute of Genetics, Fudan University, Handan Road 220, Shanghai 200433, PR China

BASE COUNT	478 a	768 c	746 g	496 t
ORIGIN				
Query Match	99.7%	Score 1070.8;	DB 9;	Length 2488;

Best Local Similarity 99.8%; Pred. No. 5.5e-188;
Matches 1072; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Oy 1 ATGCGAGCCACCCCTGCTGCTCTGCGGGTTCCTGTCGAGAGAGAGCGGTTGAG 60
Db 541 ATGCGAGCCACCCCTGCTGCTCTGCGGGTTCCTGTCGAGAGAGAGCGGTTGAG 600
Oy 61 TTGATGACAACTTATAGATACCGAGGTCCTCCGAGAAACGAGCTCCGAGGCCCCAG 120
Db 601 TTGATGACAACTTATAGATACCGAGGTCCTCCGAGAAACGAGCTCCGAGGCCCCAG 660
Oy 121 CCCAGACTGCCCCCTGCTGTTGGCCCTGAGCCCACTACTGCTCCAGATGCTCAACT 180
Db 661 CCCAGACTGCCCCCTGCTGTTGGCCCTGAGCCCACTACTGCTCCAGATGCTCAACT 720
Oy 181 GCTGTGGCACTGCTCCGCTCTTGGGCTTATGTCTCTTGGAGCCCGAGAGGCGCG 240
Db 721 GCTGTGGCACTGCTCCGCTCTTGGGCTTATGTCTCTTGGAGCCCGAGAGGCGCG 780
Oy 241 CGGGCTTACCGGGCCCTGCACTGCCCCCTACAGCACTGATATACCTGMAAGTTACCC 300
Db 781 CGGGCTTACCGGGCCCTGCACTGCCCCCTACAGCACTGATATACCTGMAAGTTACCC 840
Oy 301 GTCCAGAGAGCCCTGACGCTGTGGAGCCCTACGCGGCTGCCCCGCAACAGCATGTG 360
Db 841 GTCCAGAGAGCCCTGACGCTGTGGAGCCCTACGCGGCTGCCCCGCAACAGCATGTG 900
Oy 361 GCTGCGCCCACTAGAGTCTGCGCTGGTACCCAGCTCTCTACGCTTTTTCATCGAAC 420
Db 901 GCTGCGCCCACTAGAGTCTGCGCTGGTACCCAGCTCTCTACGCTTTTTCATCGAAC 960
Oy 421 CATGGGACATGCAACGCTGTGTGGAGCCGCGCAACCTTACCTGAGACCTGAGCTGCC 480
Db 961 CATGGGACATGCAACGCTGTGTGGAGCCGCGCAACCTTACCTGAGACCTGAGCTGCC 1020
Oy 481 GTGCTCTTCCGCGAGATGAGCCGCTGGAGCACTGTCAACAGCAAGCTGTGGCTGTG 540
Db 1021 GTGCTCTTCCGCGAGATGAGCCGCTGGAGCACTGTCAACAGCAAGCTGTGGCTGTG 1080
Oy 541 CGTATCTCAAGCTGTGTGCTTTGTCTTCCCTGACCTGAGAGAGAGAACTGGTGTG 600
Db 1081 CGTATCTCAAGCTGTGTGCTTTGTCTTCCCTGACCTGAGAGAGAGAACTGGTGTG 1140
Oy 601 GAGAACCTGAGAGACTCTGCTGTGTGCTGAGCTGGGCGAGATGATTCCTGTGGAGCAAGAG 660
Db 1141 GAGAACCTGAGAGACTCTGCTGTGTGCTGAGCTGGGCGAGATGATTCCTGTGGAGCAAGAG 1200
Oy 661 GCGTGGCCAGCTTACGTTGGAGCTGAGATTACTAGCTCAACGCGCTCTATACCTGGGCAAG 720
Db 1201 GCGTGGCCAGCTTACGTTGGAGCTGAGATTACTAGCTCAACGCGCTCTATACCTGGGCAAG 1260
Oy 721 GCAAGCCATGTCTGAGAGCTGGGCGTGGCGCTTACCATGCTGGCGGCACTACCC 780
Db 1261 GCAAGCCATGTCTGAGAGCTGGGCGTGGCGCTTACCATGCTGGCGGCACTACCC 1320
Oy 781 TTCCAGAGACTGAGAGCTGTGCTGTGGGCAAGACCGGCGGCGCTTACGCTTGG 840
Db 1321 TTCCAGAGACTGAGAGCTGTGCTGTGGGCAAGACCGGCGGCGCTTACGCTTGG 1380
Oy 841 CTTGCAAGGCTCTGCGCCCTGCGCTGTCTGTGGTCTGCTCTCTTGTGGAGGCA 900
Db 1381 CTTGCAAGGCTCTGCGCCCTGCGCTGTCTGTGGTCTGCTCTCTTGTGGAGGCA 1440
Oy 901 GCTGAGAGGCTCAACAGCAACAGGATCTCTGCAACCTTGGCTGGGCAAGAGCCGATG 960
Db 1441 GCTGAGAGGCTCAACAGCAACAGGATCTCTGCAACCTTGGCTGGGCAAGAGCCGATG 1500
Oy 961 CCCTTACGCGCAACCGGATCCCATCTGTGGAGGCTGGCGAGGTGCTCCGATGAGACTG 1020
Db 1501 CCCTTACGCGCAACCGGATCCCATCTGTGGAGGCTGGCGAGGTGCTCCGATGAGACTG 1560
Oy 1021 GGGCTGAGCAGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
```

Db 1561 GGGCTGAGCAGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1614

```
RESULT 5
AX099934
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCES
1
Yue, H., Tang, Y.T., Bandman, O., Hillman, J.L., Baughn, M.R.,
Azimzai, Y., and Lu, D.A.
Protein phosphatase and kinase proteins
Patent: WO 0120004-A 16 22-MAR-2001,
Incyte Genomics, Inc. (US)
FEATURES
source
1..2116
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID NO: 1271505CB1"
BASE COUNT
407 a 663 c 609 g 437 t
ORIGIN
Query Match 99.8%; Score 1069.2; DB 6; Length 2116;
Best Local Similarity 99.7%; Pred. No. 1.1e-187;
Matches 1071; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 1 ATGCGAGCCACCCCTGCTGCTCTGCGGGTTCCTGTCGAGAGAGAGCGGTTGAG 60
Db 161 ATGCGAGCCACCCCTGCTGCTCTGCGGGTTCCTGTCGAGAGAGAGCGGTTGAG 220
Oy 61 TTGATGACAACTTATAGATACCGAGGTCCTCCGAGAAACGAGCTCCGAGGCCCCAG 120
Db 221 TTGATGACAACTTATAGATACCGAGGTCCTCCGAGAAACGAGCTCCGAGGCCCCAG 280
Oy 121 CCCAGACTGCCCCCTGCTGTTGGCCCTGAGCCCACTACTGCTCCAGATGCTCAACT 180
Db 281 CCCAGACTGCCCCCTGCTGTTGGCCCTGAGCCCACTACTGCTCCAGATGCTCAACT 340
Oy 181 GCTGTGGCACTGCTCCGCTCTTGGGCTTATGTCTCTTGGAGCCCGAGAGGCGCG 240
Db 341 GCTGTGGCACTGCTCCGCTCTTGGGCTTATGTCTCTTGGAGCCCGAGAGGCGCG 400
Oy 241 CGGGCTTACCGGGCCCTGCACTGCCCCCTACAGCACTGATATACCTGMAAGTTACCC 300
Db 401 CGGGCTTACCGGGCCCTGCACTGCCCCCTACAGCACTGATATACCTGMAAGTTACCC 460
Oy 301 GTCCAGAGAGCCCTGAGAGCTGTGCTGTGGGCAAGACCGGCGGCGCTTACGCTTGG 360
Db 461 GTCCAGAGAGCCCTGAGAGCTGTGCTGTGGGCAAGACCGGCGGCGCTTACGCTTGG 520
Oy 361 GCTGCGCCCACTAGAGTCTGCTGGTGTATACCGAGCTCTTACGCTTTTTCATCGAAC 420
Db 521 GCTGCGCCCACTAGAGTCTGCTGGTGTATACCGAGCTCTTACGCTTTTTCATCGAAC 580
Oy 421 CATGGGACATGCAACGCTGTGTGGAGCCGCGCAACCTTATCCCTGAGACCTGAGCTGCC 480
Db 581 CATGGGACATGCAACGCTGTGTGGAGCCGCGCAACCTTATCCCTGAGACCTGAGCTGCC 640
Oy 481 GTGCTCTTCCGCGAGATGAGCCGCTGGGCGACTGTGTCAACAGCAAGCTGTGGCTGTG 540
Db 641 GTGCTCTTCCGCGAGATGAGCCGCTGGGCGACTGTGTCAACAGCAAGCTGTGGCTGTG 700
Oy 541 CGTATCTCAAGCTGTGTGCTTTGTCTTGTGCTGACCGTGTGAGAGAGAGAGAGAGAG 600
Db 701 CGTATCTCAAGCTGTGTGCTTTGTCTTGTGCTGACCGTGTGAGAGAGAGAGAGAGAG 760
```


QY 781 TTCCAGAGCTCGAGGCTGTCTGTCTTTCGAGAGATCGCGCGGAGCTTACGCGCTTG 840
DB 1075 TTCCAGAGCTCGAGGCTGTCTGTCTTTCGAGAGATCGCGCGGAGCTTACGCGCTTG 1134
QY 841 CCTGAGGCTCTCGGCGGCTGCGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 900
DB 1135 CCTGAGGCTCTCGGCGGCTGCGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1194
QY 901 GCTGAACGCTGACAGAGCAAGGAGATCTCTGAGAGATCGCGCGGAGCTTACGCGCTTG 960
DB 1195 GCTGAACGCTGACAGAGCAAGGAGATCTCTGAGAGATCGCGCGGAGCTTACGCGCTTG 1254
QY 961 CCTTACGCGCAACCGGATCCCATCTCTGAGAGAGCTGCGCGGAGCTTACGCGCTTG 1020
DB 1255 CCTTACGCGCAACCGGATCCCATCTCTGAGAGAGCTGCGCGGAGCTTACGCGCTTG 1314
QY 1021 GGGCTGAG 1074
DB 1315 GGGCTGAG 1368

RESULT 7
AX572896 1083 bp DNA linear PAT 29-NOV-2002
LOCUS Sequence 3 from Patent WO02053743.
DEFINITION AX572896
ACCESSION AX572896
VERSION AX572896.1 GI:26004968
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Dower, S. and Quarnstrom, B.
Mammalian triblbles signaling pathways and methods and reagents
related thereto
Patent: WO 02053743-A 3 11-JUL-2002;
JOURNAL Interleukin Genetics, Inc. (US)
FEATURES
Location/Qualifiers
source 1..1083
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 176 a 374 c 327 g 206 t
ORIGIN

Query Match 97.9%; Score 1051.6; DB 6; Length 1083;
Best Local Similarity 99.1%; Pred. No. 2,3e-184;
Matches 1070; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 1 ATGCGAGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
DB 1 ATGCGAGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 61 TTGATGACAACTTATGATATCGAGCGTCCGCTCGAGAAAGAGCTGAGAGAGAGAGAG 120
DB 61 TTGATGACAACTTATGATATCGAGCGTCCGCTCGAGAAAGAGCTGAGAGAGAGAGAG 120
QY 121 CCGAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 121 CCGAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 181 GCTGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 181 GCTGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 CCGAGCTACCGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 241 CCGAGCTACCGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 301 GTCCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 GTCCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

DB 301 GTCCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 GCTGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GCTGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 CATGGGAGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 CATGGGAGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GCTGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GCTGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CCGATCTGAG 594
DB 541 CCGATCTGAG 600
QY 595 GTGCTGAG 654
DB 601 GTGCTGAG 660
QY 655 AAGCAGCGGTGCCAGAGCTTACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
DB 661 AAGCAGCGGTGCCAGAGCTTACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 715 GGCAG 774
DB 721 GGCAG 780
QY 775 TACCCCTTCAG 834
DB 781 TACCCCTTCAG 840
QY 835 GCTTTCCTGAG 894
DB 841 GCTTTCCTGAG 900
QY 895 GAGCAG 954
DB 901 GAGCAG 960
QY 955 CCGATGCTTACGCGCAACCGGATCCCATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
DB 961 CCGATGCTTACGCGCAACCGGATCCCATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1015 GGAAG 1074
DB 1021 GGAAG 1080

RESULT 8
AF250311 1083 bp mRNA linear PRI 01-JUN-2001
LOCUS Homo sapiens SKIP3 mRNA, complete cds.
DEFINITION AF250311
ACCESSION AF250311
VERSION AF250311.1 GI:14276268
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1083)
Kies-Toth, B., Wyllie, D.H., Quarnstrom, B.B. and Dower, S.K.
Identification of pro-inflammatory cytokine signaling network
components by transcription expression screening
Unpublished
2 (bases 1 to 1083)
Kies-Toth, B., Wyllie, D.H., Quarnstrom, B.B. and Dower, S.K.
Direct Submission
Submitted (29-MAR-2000) Division of Molecular & Genetic Medicine,
University of Sheffield, Royal Hallamshire Hospital, Floor M,
Sheffield S10-2UF, UK

Db 241 CGGCGCTACAGAGCCCTGCACTGGCTTACAGGCACTGAGTATACCTGAGAGGTATACCC 300
Qy 301 GTCCAGAGAGCCCTGCGCTGCTGAGAGCCCTACGCGCGCTGCCCCGACAGAGATGTG 360
Db 301 GTCCAGAGAGCCCTGCGCTGCTGAGAGCCCTATGCGCGCTGCCCCGACAGAGATGTG 360
Qy 361 GCTGGGCGCCACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 GCTGGGCGCCACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 421 CATGGGAGCATGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 CATGGGAGCATGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 446
Qy 481 GTGCTCTTCCGCGAGATGCGCCCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 447 -----GCACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
Qy 541 CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 475 -----ACAGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
Qy 601 GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 496 GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555
Qy 661 GCGTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 556 GCGTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
Qy 721 GCAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 616 GCAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 675
Qy 781 TTCCAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 676 TTCCAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 735
Qy 841 CTTGCAAGGCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 736 CTTGCAAGGCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
Qy 901 GCTGAGACGCTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 796 GCTGAGACGCTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 855
Qy 961 CCCTTAGGCGCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 856 CCCTTAGGCGCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 915
Qy 1021 GGGCTGAG 1074
Db 916 GGGCTGAG 969

RESULT 10
AF358868
LOCUS AF358868 1256 bp mRNA linear ROD 01-JUN-2002
DEFINITION Mus musculus TRB-3 mRNA, complete cds.
ACCESSION AF358868.1 GI:21304713
VERSION AF358868.1 GI:21304713
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1256)
Kies-Toth,B., Dempsey,C., Jozsa,V., Caunt,J., Oxley,K.M.,
Bagatsis,S.M., Myllye,D.H., Harter,M., O'Neill,L.A.J.,
Ovarietrom,B.E. and Dower,S.K.
Mammalian homologs of Drosophila tritubules (httb) control mitogen
activated protein kinase signaling

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1256)
AUTHORS Kies-Toth,B., Dempsey,C., Jozsa,V., Caunt,J., Oxley,K.M.,
Bagatsis,S.M., Myllye,D.H., Harter,M., O'Neill,L.A.J.,
Ovarietrom,B.E. and Dower,S.K.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2001) Division of Genomic Medicine, University of
Sheffield, Royal Hallamshire Hospital, Floor M, Glossop Road,
Sheffield S10-2JF, UK
FEATURES
source
1..1256
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
182..1246
/note="trb-3"
/codon_start=1
/product="TRB-3"
/protein_id="AA045476.1"
/db_xref="gi:21304713"
/translation="MRAPTLAASADVGRKPLSPDNDIAKCPVLKVRDPRNCP
PSLPSPPSPASDLSPAVAFATRIAPITLIRBQSGSYRALHCTGTBCTKTPAS
EAQVLAIPYARLPFHQVAPTEVLDSRLLYIFPKTHGDLHSLSRSRRI
GLPRQMASAVAHCHKGIVLRLRPFVFNCRITKLVLRLBDACTMTGSDSLMD
THACPAYGPEILSRPSYSGKADVMSLGVALLFTMLAGRPVPHDSVPLFGKIRXG
TPALPGLSAPARLCTIRLKEPSRLVALGILHPVLRDHDHGVSPQSDRBMDO
VVPDGPOLBRLBBERVGLXG"
BASE COUNT 231 a 390 c 385 g 249 t 1 others
ORIGIN
Query Match 61.1%; Score 656.6; DB 10; Length 1256;
Best Local Similarity 76.6%; Pred. No. 2e-111;
Matches 803; Conservative 0; Mismatches 245; Indels 0; Gaps 0;
Qy 1 ATGCGAGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 182 ATGCGAGCTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
Qy 61 TTGATATGCACTTATGATACCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 242 TTGATATGCACTTATGATACCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
Qy 121 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 302 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
Qy 362 GCTGTGGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 421
Db 422 TGCACTATGAGAGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 481
Qy 481 GCTGTGGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 540
Db 542 GCTGTGGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 601
Qy 601 GCTGTGGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 661
Db 662 GCTGTGGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 721
Qy 721 GCTGTGGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 781
Db 782 GCTGTGGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 841

Qy 781 TTCAGAGACTGGAGCCTGTCTCTCTTCCGCAAGATCCGCCGCCCTTACGCCCTTG 840
Db 919 TTCAGAGACTGTGAACAGCCCTCTCTTGGCAAGATCCGTAGAGGACCTTTGCCCTG 978
Qy 841 CCTCAGAGCCTCTCGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 979 CTTAGAGGCTTATAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
Qy 901 GCTTAAACGGCTCAGACGCCAGCAGCATCTCTGCAACCCCTGCTGCTGCTGCTGCTGCTG 960
Db 1039 TCAGAGGAGACTTGTGGCCCTGGGAATCTGCTACATCTCTGTTGAGGAGAGATTGGACG 1098
Qy 961 CCTTACCCCAACCCCAATCCATCTGCTGAGAGGCTGCGAGGCTGCTGCTGCTGCTGCTG 1020
Db 1099 CAACTCTCTCTCAACGCTGCTGACCGAAGGAGATGAGCAAGTTGTCTCCAGATGGGCCA 1158
Qy 1021 GAGCTGAGACGAGAGCCAGAGAGAGAGAG 1048
Db 1159 CAGCTGAGAGAGGAGAGAGAGAGAGAG 1186

RESULT 12

AX364906 1076 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 57 from Patent WO0206315.
DEFINITION AX364906
ACCESSION AX364906
VERSION AX364906.1 GI:18696795
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Mintz, L., Freilich, S. and Bernstein, J.
TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 57 24-JAN-2002;
Compugen Ltd. (IL)

FEATURES
Source Location/Qualifiers
1..1076
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 172 a 352 c 346 g 194 t 12 others
ORIGIN

Query Match 55.6%; Score 597.2; DB 6; Length 1076;
Best Local Similarity 98.1%; Pred. No. 1.9e-100;
Matches 622; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

Qy 1 ATGGAGCCACCCCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 329 ATGGAGCCACCCCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 388
Qy 61 TTGGATGACAACTTAGATACCGAGCGTCCGTCAGAAACGAGCTCGAAGTGGGCCCGAG 120
Db 389 TTGGATGACAACTTAGATACCGAGCGTCCGTCAGAAACGAGCTCGAAGTGGGCCCGAG 448
Qy 121 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 449 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 508
Qy 181 GCTGTGACCACTGCTCCCGCTTGGGGCCCTATGCTCCCTGAGAGCCCGGAGAGAGAGCGGG 240
Db 509 GCTGTGACCACTGCTCCCGCTTGGGGCCCTATGCTCCCTGAGAGCCCGGAGAGAGAGCGGG 568
Qy 241 CGGGCTTACCGGGCCCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 569 CGGGCTTACCGGGCCCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
Qy 301 GTCCAGAGAGCCCTGCGCGCTGCTGAGAGCTTACGCGGAGCTGCGCGGAGAGAGATG 360
Db 629 GTCCAGAGAGCCCTGCGCGCTGCTGAGAGCTTACGCGGAGCTGCGCGGAGAGAGATG 688

Qy 361 GCTGAGCCCACTAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 689 GCTGAGCCCACTAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
Qy 421 CATGGAGACATGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 749 CATGGAGACATGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
Qy 481 GTGCTCTTCCGCGCAGATGCGCAACCGCTGCGCACTGTCAACAGCAAGTGTGCTGCTG 540
Db 809 GTGCTCTTCCGCGCAGATGCGCAACCGCTGCGCACTGTCAACAGCAAGTGTGCTGCTGCTG 868
Qy 541 CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 869 CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927
Qy 601 GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 634
Db 928 GAGA--CCTGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959

RESULT 13

AX364921 1076 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 72 from Patent WO0206315.
DEFINITION AX364921
ACCESSION AX364921
VERSION AX364921.1 GI:18696810
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Mintz, L., Freilich, S. and Bernstein, J.
TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 72 24-JUN-2002;
Compugen Ltd. (IL)

FEATURES
Source Location/Qualifiers
1..1076
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 172 a 352 c 346 g 194 t 12 others
ORIGIN

Query Match 55.6%; Score 597.2; DB 6; Length 1076;
Best Local Similarity 98.1%; Pred. No. 1.9e-100;
Matches 622; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

Qy 1 ATGGAGCCACCCCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 329 ATGGAGCCACCCCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 388
Qy 61 TTGGATGACAACTTAGATACCGAGCGTCCGTCAGAAACGAGCTCGAAGTGGGCCCGAG 120
Db 389 TTGGATGACAACTTAGATACCGAGCGTCCGTCAGAAACGAGCTCGAAGTGGGCCCGAG 448
Qy 121 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 449 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 508
Qy 181 GCTGTGACCACTGCTCCCGCTTGGGGCCCTATGCTCCCTGAGAGCCCGGAGAGAGAGCGGG 240
Db 509 GCTGTGACCACTGCTCCCGCTTGGGGCCCTATGCTCCCTGAGAGCCCGGAGAGAGAGCGGG 568
Qy 241 CGGGCTTACCGGGCCCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 569 CGGGCTTACCGGGCCCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
Qy 301 GTCCAGAGAGCCCTGCGCGCTGCTGAGAGCTTACGCGGAGCTGCGCGGAGAGAGATG 360
Db 629 GTCCAGAGAGCCCTGCGCGCTGCTGAGAGCTTACGCGGAGCTGCGCGGAGAGAGATG 688


```

Oy 361 GCTGGGCCCCATGAGTCTGTGGTGTACCAAGCTCTACGCGCTTTTCACTGGACC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 GCCGTCCCAAGAGGTCCTGTGGCTCTGGCTCTTTTACATTTTTCACAAAGACC 596
Oy 421 CATGGGACATGACAGCGCTGTGGAGAGCCGCCACCGTATCCCTGAGCCGTGAGTCC 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 597 CATGGGACATGACAGCGCTGTGGAGAGCCGCCACCGTATCCCTGAGCCGTGAGTCC 656
Oy 481 GTGCTTTCCGCCAGATAGCCAGCCGCTGTGGAGAGTCTGACAGAGCGTGTGCTCTG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 657 GGGCTCTTCCGAGAGATGCTAGTGCCTGGACACTGCACACAGACAGCGCTGTCTTG 716
Oy 541 CGTATCTCAAGCTGTGTGCTTTTGTCTTCCCTGACCGTGAAGAGAAAGCTGGTCTG 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 717 GCGACCTCAAGCTGTGTGCTTTTGTCTTCAAGCAACTGTGAGAGAGCAAGCTGTCTG 776
Oy 601 GAGAACTGAGAGACTCTGTGGTGTGACTGTGGCCAGATATCTCTGTGGAGCAAGAC 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 777 GAGAACTGAGAGACTCTGTGGTGTGACTGTGGCCAGATATCTCTGTGGAGCAAGAT 836
Oy 661 GCGTGGCCAGCTTACGTGGAGACTGTGATCTCAAGCTCAAGCGCTCATCTCGGGCAG 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 837 GCGTGGCCAGCTTACGTGGAGACTGTGATCTCAAGCTCAAGCGCTCATCTCGGGCAG 885
Oy 721 GCAAGCCAGATGTGTGAGAGCTGTGGCGTGTGCTTCAACATGTGTGGCCAGCTAACC 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 ----- 885
Oy 781 TTCAGAGACTGAGAGCTGTGTCTGTCTTGGGAGATCGCCCGGGGCGCTACGCTTG 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 -----ACTGTGAGCCAGCTCTGTCTTGTGGCAAGATCCGTAGAGGAGACCTTGGCCCTG 938
Oy 841 CTTGACAGGCTCTGTGGGCGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 939 CTTGAGGCGCTTATAGAGCCGAGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998
Oy 901 GCTGAACGGCTCACAGCCAGAGGATCTCTGTGCAAGCCGCTGTGTGTGTGTGTGTGTGT 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 999 TCAAGAGGACTTGTGGGCGCCGAGATCTTGTGCAATCCCTGTGTGTGTGTGTGTGTGT 1058
Oy 961 CCTTAAAGCCCAAGCCGATCTCTGTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGT 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1059 CGAGTCTCTCTCTCAAGCTGTGACCGAGAGGAGATGACACAGTGTGTGTGTGTGTGT 1118
Oy 1021 GGGCTGAGAGAGCGAGGAGAGAGAG 1048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1119 CAGCTGAGAGAGCGTGAAGAGAGAG 1146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

LOCUS HS1103G7 153170 bp DNA linear PRI 08-FEB-2001

DEFINITION Human DNA sequence from clone RPS-1103G7 on chromosome 20p12.2-13. Contains up to three novel genes, the gene for a novel protein similar to mouse VMP, the gene for a novel protein kinase domain containing protein similar to phosphoprotein C8PW and rat NPK, and the SOX22 gene for SRY (sex-determining region Y)-box 22. Contains five CpG islands, ESTs, STSs and GSSs, complete sequence.

ACCESSION AL034548 GI:7263904

VERSION 1.0

KEYWORDS HMG; CpG island; NPK; protein kinase; SOX22; SRY; VMP.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 153170)

AUTHORS Blakey, S.

TITLE Direct Submision

JOURNAL Submitted (05-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, UK

COMMENT On Mar 13, 2000 this sequence version replaced gi:5541861. During sequence assembly data is compared from overlapping clones.

FEATURES

Source

Location/Qualifiers

1..153170

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="20"

/map="p12.2-13"

/clone="RPS-1103G7"

/clone_id="RPS1-5"

1..61

/note="Alu repeat: matches 2. .62 of consensus"

79..129

/note="L1M9 repeat: matches 6221. .6271 of consensus"

130..173

/note="22 copies 2 mer ag 79% conserved"

175..873

/note="L1M9 repeat: matches 5519. .6217 of consensus"

902..1109

/note="L1M5 repeat: matches 7715. .7921 of consensus"

1110..1480

/note="THB1C repeat: matches 1. .371 of consensus"

1502..1733

/note="match: STS: Bm:26755"

1734..1888

/note="LTR repeat: matches 387. .547 of consensus"

1919..2101

/note="LTR repeat: matches 115. .298 of consensus"

complement(join(2146..2243,126792..126834))

/note="match: GSS: Bm:B59651"

2154..2193

/note="20 copies 2 mer aa 85% conserved"

complement(join(2194..2243,126792..126834))

/note="match: GSS: Bm:AQ37604"

2699..2917

/note="L1M4 repeat: matches 2675. .3081 of consensus"

2918..3062

/note="L1M4 repeat: matches 3120. .3249 of consensus"

3136..3430

/note="L1M9 repeat: matches 1. .295 of consensus"

3462..3777

/note="L1M9 repeat: matches 1. .296 of consensus"

3811..4126

/note="L1M9 repeat: matches 1. .312 of consensus"

4310..4608

/note="L1M2 repeat: matches 5861. .6167 of consensus"

4609..5026

/note="MSRA repeat: matches 1. .424 of consensus"

5027..5200

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Bm, BMB; Sw, SWISSPROT; Tr, TREMBL; Mp, MORNREP; Information on the MORNREP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormrep This sequence is the entire insert of clone RPS-1103G7 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

RPS-1103G7 is from the library RPS1-5 constructed by the group of Pletzer de Jong. For further details see <http://www.choi.org/bacpac/home.htm>

VECTOR: pCYPAC2.

Search completed: January 16, 2004, 14:49:08
Job time : 4032.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 10:03:56 ; Search time 4616.16 Seconds
(without alignments)
10840.821 Million cell updates/sec

Title: US-09-909-474D-1

Perfect score: 2059 1 gctctgagcccgccgagcgc.....aaaaaaaaaaaaaaaa 2059

Sequence: 1 gctctgagcccgccgagcgc.....aaaaaaaaaaaaaaaa 2059

Scoring table: IDENTITY_MUC Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_gatba:*
2: em_gatba:*
3: em_gatba:*
4: em_gatba:*
5: em_gatba:*
6: em_gatba:*
7: em_gatba:*
8: em_gatba:*
9: gb_gatba:*
10: gb_gatba:*
11: gb_gatba:*
12: gb_gatba:*
13: gb_gatba:*
14: gb_gatba:*
15: em_gatba:*
16: em_gatba:*
17: em_gatba:*
18: em_gatba:*
19: em_gatba:*
20: em_gatba:*
21: em_gatba:*
22: em_gatba:*
23: em_gatba:*
24: em_gatba:*
25: em_gatba:*
26: em_gatba:*
27: em_gatba:*
28: gb_gatba:*
29: gb_gatba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	940.6	45.7	1201	9	AL578892
C 2	902	43.8	942	13	BX393466
C 3	895.2	43.5	1075	13	BX367264
C 4	893	43.4	1051	9	AL525890

C 5	892	43.3	947	13	BX325344
C 6	868.8	41.9	975	10	BG575275
C 7	862.4	42.2	1001	9	AL562576
C 8	849.4	41.1	873	13	BQ216198
C 9	846.2	41.1	881	13	BQ941789
C 10	839.8	40.8	945	13	BX325345
C 11	834	40.5	1201	13	BX426485
C 12	832.8	40.4	1201	13	BX445733
C 13	818.6	39.8	863	10	BG748142
C 14	818.4	39.7	1201	9	AL556690
C 15	816.6	39.6	948	13	BX443303
C 16	814.6	39.7	902	13	BX421477
C 17	811.2	39.4	1062	13	BM557568
C 18	807.2	39.2	837	12	B1861756
C 19	806.2	39.2	918	13	BQ223457
C 20	802.4	39.0	1036	13	BQ217691
C 21	797.8	38.7	925	13	BX435881
C 22	796.6	38.7	946	13	BX390085
C 23	790.2	38.4	929	13	BUS00619
C 24	788	38.3	929	9	AL522632
C 25	765.6	37.2	1126	13	BX405957
C 26	759	36.9	944	13	BX363066
C 27	756	36.7	914	13	BUS00176
C 28	753.8	36.6	1076	13	BX367265
C 29	744.6	36.2	851	12	B1909168
C 30	744	36.1	820	12	B1256421
C 31	740.6	36.0	880	10	BG470123
C 32	738.6	35.9	1002	12	BM804790
C 33	737	35.8	910	9	AL562114
C 34	736.4	35.8	892	14	CA495347
C 35	735.8	35.7	1053	13	BQ224844
C 36	735.2	35.7	861	9	AL555266
C 37	735	35.7	989	10	BE792929
C 38	730.2	35.5	1051	12	BM907602
C 39	728.2	35.4	810	12	BM045379
C 40	725.4	35.2	838	10	BG171687
C 41	718	34.9	792	10	BG387820
C 42	714.6	34.7	931	10	BG470193
C 43	713.2	34.6	957	13	BQ949069
C 44	712.8	34.6	918	9	AL520575
C 45	709.4	34.5	910	10	BG397691

ALIGNMENTS

RESULT 1
AL578892/c 1201 bp mRNA linear EST 01-JUN-2003
DEFINITION AL578892 Homo sapiens HEPA CELLS COT 25-NORMALIZED Homo sapiens
LOCUS CDNA Clone CS0DK005YC05 3-PRIME, mRNA sequence.
ACCESSION AL578892
VERSION AL578892.2 GI:31317080
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS I.A.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced gi:12943405.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10000.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK005YAB03NP1&cluster=10000.f. Contact :
Peng Liang Email : liang@lifetech.com URL :

QY	1155	CAGTGGATTGAACTTTGGGGGTAGCTTCAAGACCTTCTCCGCTCTGAAGTGAACCAACC	121
Db	822	CAGTGGATTGAACTTTGGGGGTAGCTTCAAGACCTTCTCCGCTCTGAAGTGAACCAACC	763
QY	1215	TTCAAGTGCCTTTCAGAAGGGAGAAAGGAGAAAGCCTGTGTGGAGTGTGCTGTGACACT	1274
Db	762	TTCAAGTGCCTTTCAGAAGGGAGAAAGGAGAAAGCCTGTGTGGAGTGTGCTGTGACACT	703
QY	1275	CTGCTTTTGTCCACACATGCAATGCAATTCCTGTGGGTGCTTATCAAGTGGCCAAAGCCGTG	1333
Db	702	CTGCTTTTGTCCACACATTTTBAAGTTCTGTGTTGGGTGCTTATCAAGTGGCCAAAGCCGTG	643
QY	1335	TCTCGGTCTGGAGTATCAGCAGTGAAGCAAGAGAGACAATATTCCTGTCTCAAGAGATG	1394
Db	642	TCTCGGTCTGGAGTATCAGCAGTGAAGCAAGAGAGACAATATTCCTGTCTCAAGAGATG	583
QY	1355	ACAAATCTGGCATCTTGTAGCTGACAAACTTTTCCATATACATTAAGTCACTGTCTTACACT	1455
Db	582	ACAAATCTGGCATCTTGTAGCTGACAAACTTTTCCATATACATTAAGTCACTGTCTTACACT	523
QY	1455	GGGTACACTTTTGTACAGTGTGGGCTCCACTGATAGTGGTGTCTCAGGACACTGTGTCCA	1511
Db	522	GGGTACACTTTTGTACAGTGTGGGCTCCACTGATAGTGGTGTCTCAGGACACTGTGTCCA	463
QY	1515	AGGACAATTCCTTTTCAAAACAAACCAAGCTGCTTTTGTATCTTGTACCTTTTCAAGAA	1574
Db	462	AGGACAATTCCTTTTCAAAACAAACCAAGCTGCTTTTGTATCTTGTACCTTTTCAAGAA	403
QY	1575	GGGAGGTATCCCTGTGGCCAAAGGCTCGAGGCTCTCCCTGTGCACTCAGGACCCAAAGCC	1633
Db	402	GGGAGGTATCCCTGTGGCCAAAGGCTCTGAGGCTCTCCCTGTGCACTCAGGACCCAAAGCC	343
QY	1635	AGCTCACTCTGGGAACTGTGTTCCGACGACTCTGTGCTCTTGAATTAAGAGATTCTCCTT	1694
Db	342	AGCTCACTCTGGGAACTGTGTTCCGACGACTCTGTGCTCTTGAATTAAGAGATTCTCCTT	283
QY	1655	CCAGGCTCTAAGCTCTGGGATTTGGGCGACAGATTAATAATCCAACTATAGGCTAATTCTT	1755
Db	282	CCAGGCTCTAAGCTCTGGGATTTGGGCGACAGATTAATAATCCAACTATAGGCTAATTCTT	223
QY	1755	GTCTAATCTCAAGACTGTTCTGGAATGAGGGTCTCAGGCTGTGCAACATGAGGCTTCTGAC	1814
Db	222	GTCTAATCTCAAGACTGTTCTGGAATGAGGGTCTCAGGCTGTGCAACATGAGGCTTCTGAC	163
QY	1815	CTGAGCACCAAGGTTGAGGAGACAGATTAAGGCAAGGCTGTGCTGTGGCCACTCTGGAAG	1874
Db	162	CTGAGCACCAAGGTTGAGGAGACAGATTAAGGCAAGGCTGTGCTGTGGCCACTCTGGAAG	103
QY	1875	TCCGAGGTGGGACTCTCTTCGGGAGACACTTGGGGTCCAAATCCGAGGTCCATCTTAGG	1933
Db	102	CCCGAGGTGGGACTCTCTTCGGGAGACACTTGGGGTCCAAATCCGAGGTCCATCTTAGG	43
QY	1935	TTTTGGATACATGAGTATGTATGTTTACTGTG	1968
Db	42	YCCTNNANACGGAGATTATTTATTTTACTTACB	9
RESULT 3			
EX367264/c			
LOCUS	EX367264	1075 bp	mRNA linear EST 08-MAY-2003
DEFINITION	EX367264 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED		
ACCESSION	Homo sapiens cDNA clone C8DD002Y07 3-PRIME, mRNA sequence.		
VERSION	EX367264		
KEYWORDS	EX367264.1 GI:30441034		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Li, W.B., Gruber, C., Jeejee, J. and Polayes, D.		
	Full-length cDNA libraries and normalization		
	Unpublished		

COMMENT	Contact : Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email : seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10000.f. For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0AL002CB04NP1&cluster=10000.f . Contact : Peng Liang Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com/ , Invitrogen Corporation 1600 Paradey Avenue Genoscope sequence ID : CS0AL002CB04NP1. location/Qualifiers
FEATURES	source
	1..1075
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CS0DL002J07"
	/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
	/cell_line="RAMOS CELL LINE"
	/clone_id="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
	/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASIS COUNT	264 a 256 c 281 g 243 t 31 others
ORIGIN	
Query Match	43.5%; Score 895.2; DB 13; Length 1075;
Best Local Similarity	95.7%; Pred. No. 6.7e-73;
Matches 929,	Conservative 10; Mismatches 30; Indels 2; Gaps 2;
OY	AGCCCAACCGATGCCCATCTGTGGAGGCTGCCAAGTGTCCTTAGTGAATGAGCTGGGGCT 1073
DB	ATCCTTAATCAACCACMCATTTTGGAGACTCCCAAGTGCTCTATGAACTGGGGCT 911
OY	GAGCAAGCCAGGAGAAGAGAGAGACAAGAAAGTGTTCTGATGAGCTAACCAACC 1133
DB	KGAAGARCCAGGGAMAGGA-GGARACAGAGAAKTGGTTCTATGAGCTARACACCC 852
OY	TACTACAGCGCTAGCTGCCAAGTGAATTGAGTTGGGGTACTCCAAGCTTCTCT 1199
OY	TACTACAGCGCTAGCTGCCAAGTGAATTGAGTTGGGGTACTCCAAGCTTCTCT 792
DB	851 TACTACAGCGCTAGCTGCCAAGTGAATTGAGTTGGGGTACTCCAAGCTTCTCT 792
OY	GCCTCTGAAGTGAAGCCAAACCTTCAGTGCCTTCAGAAAGGAGAAAGCAAGCTGTG 1253
DB	791 GCCTCTGAAGTGAAGCCAAACCTTCAGTGCCTTCAGAAAGGAGAAAGCAAGCTGTG 732
OY	TGAAGTGTCTGTGTACATCTGCTTTGTTCCACAATGACATGTCAGTTCTGTGGGTGC 1313
DB	731 TGAAGTGTCTGTGTACATCTGCTTTGTTCCACAATGACATGTCAGTTCTGTGGGTGC 672
OY	TTATCAGATGCAAGCCCTGTTCTCGGTGCTGGAGTACAGCAATGCAAGAGACAA 1373
OY	TTATCAGATGCAAGCCCTGTTCTCGGTGCTGGAGTACAGCAATGCAAGAGACAA 612
DB	671 TTATCAGATGCAAGCCCTGTTCTCGGTGCTGGAGTACAGCAATGCAAGAGACAA 612
OY	TATTCCTGTCTCACAGAGATGACAAATGCGATCTTGAAGTGAACAATTTTCAATGA 1433
DB	611 TATTCCTGTCTCACAGAGATGACAAATGCGATCTTGAAGTGAACAATTTTCAATGA 552
OY	CCATAGGTCACTGTCTACATCTGGGTACATTTTGAACAGATGTGGGCTTCCATGATGTG 1493
DB	551 CCATAGGTCACTGTCTACATCTGGGTACATTTTGAACAGATGTGGGCTTCCATGATGTG 492
OY	GTGCTCAGGACCTCTGTCCAAAGCAATCCCTTTTCAAACCAACAGCTGCTTTGTGA 1553
DB	491 GTGCTCAGGACCTCTGTCCAAAGCAATCCCTTTTCAAACCAACAGCTGCTTTGTGA 432
OY	TCTTGTACCTTTTCAAGAAAGGAGATATCCCTGTGCAAAAGCTTCAGAGCTCTCCCC 1613
DB	431 TCTTGTACCTTTTCAAGAAAGGAGATATCCCTGTGCAAAAGCTTCAGAGCTCTCCCC 372

Oy		1614	TGCAACTAGACCCGAACCAGCCTCACCTCGGGAACTGTGTTCACATCTCTGCCT	1673
Dd		371	TGCACTAGAGACCCAAGCCCAGCTCATCTCGGAACTGTGTTCACATCTCTGCCT	312
Oy		1674	CTTGATTAGAAGATTCTCTTCCAGGCCCTTAAGCTTTGGCCAGAGATAAAGATC	1733
Dd		311	CTTGATTAGAAGATTCTCTTCCAGGCCCTTAAGCTTTGGCCAGAGATAAAGATC	252
Oy		1794	GTCACAACATYGGGGCTTCTGACCTGACGACCAAAGTTTAGAGGACAGATTAAGCAGGGCT	1853
Dd		191	GTCAACCATYGGGGCTTCTGACCTGACGACCAAAGTTTAGAGGACAGATTAAGCAGGGCT	132
Oy		1854	GTCCTGTGGCCACCTGGAAGTCCGAGGTGGAGCTCTTCTGGGACACATTGGGGTCCACA	1913
Dd		131	GTCCTGTGGCCACCTGGAAGTCCGAGGTGGAGCTCTTCTGGGACACATTGGGGTCCACA	72
Oy		1914	ATCCCAAGTCCATACTCTTAGAGTTTTGGATTACC-ATGAATATATATATTACTTGTGCTTA	1972
Dd		71	ATCCCAAGTCCATACTCTTAGAGTTTTGGATTACCMATGATATATATTACTTGTBYTA	12
Oy		1973	ATAAAAGAGAA	1983
Dd		11	ATAAAAGAGAA	1
RESULT_4 AL525890/c LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL COMMENT		AL525890 AL525890 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo saplens CDNA clone CSODC013YD08 3-P-RIME, mRNA sequence. AL525890 AL525890.2 GI:31063754 EST. Homo saplens (human) Homo saplens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Li M.B., Gruber,C., Jesssee,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished On Feb 13, 2001 this sequence version replaced gi:12789383.	1051 bp mRNA linear .EST 23-MAY-2003 NEUROBLASTOMA COT 25-NORMALIZED Homo saplens CDNA clone CSODC013YD08 3-P-RIME, mRNA sequence. GI:31063754 EST. Homo saplens (human)	1673 312 1733 252 1853 132 1913 72 1972 12 1983 1
FEATURES SOURCE		Contact : Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seque@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequene cluster 10000.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODC013DB04NP&icuster=10000.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ invitrogen Corporation 1600 Parady Avenue Genoscope secume ID : CSODC013DB04NP. Location/Oallifiers 1..1051 /organism=""Homo saplens"" /mol_type=""mRNA"" /db_xref=""taxon:9606"" /clone=""CSODC013YD08"" /tissue_type=""NEUROBLASTOMA COT 25-NORMALIZED"" /cclone_lib=""Homo saplens NEUROBLASTOMA COT 25-NORMALIZED"" /note=""First strand cdna was primed with a NotI-oilgo(drr) primer. Five prime end enriched, double-strand cdna was digested with Not I and cloned into the Not I and EcoR V sites of the pcwmSPORT 6 vector. Library was normalized.""	1051 bp mRNA linear .EST 23-MAY-2003 NEUROBLASTOMA COT 25-NORMALIZED Homo saplens CDNA clone CSODC013YD08 3-P-RIME, mRNA sequence. GI:31063754 EST. Homo saplens (human) Homo saplens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Li M.B., Gruber,C., Jesssee,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished On Feb 13, 2001 this sequence version replaced gi:12789383.	1673 312 1733 252 1853 132 1913 72 1972 12 1983 1
J3 COUNT ORIGIN		240 a 259 c 278 g 233 t 41 others		

Query Match	43.4%	Score 893	DB 93	Length 1051
Best Local Similarity	95.7%	Pred. No. 1,1e-72		
Matches 908	Conservative 25	Mismatches 14	Indels 2	Gaps 2
Qy 1010	CCTTAGGCCCAACC	CGATCCCATCTCTGGAGGCTGCCAGCTGGTCCCTGATGACCTGG	1065	
Db 949	CCTTAGGCTCCMCCCGATCCCATCTCTGGAGGCTGCCAGGTCGCTCATATGACCTGG	890		
Qy 1070	GGCTGGACGAAGCCAGAGGAAAGAGAGGAGACAGAGAAATGGTCTCTGATATGGCTAGAAC	1122		
Db 889	GGCTGGACGAAGCCAGAGGAAAGAGAGGAGACAGAGAAATGGTCTCTGATATGGCTAGAAC	830		
Qy 1130	ACCCCTACTACACGCTCAGCTGACCCAAACAGTGAATGAGTTGGGGGTAGCTCCAAAGCTTC	1188		
Db 829	ACCCCTACTACACGCTCAGCTGACCCAAACAGTGAATGAGTTGGGGGTAGCTCCAAAGCTTC	770		
Qy 1190	TCTGCTCTCTGAACTGAGCCAAACCTTCAGTGCCTTCAGAAAGGAGAAAGGACAGAAC	1245		
Db 769	TCTGCTCTCTGAACTGAGCCAAACCTTCAGTGCCTTCAGAAAGGAGAAAGGACAGAAC	710		
Qy 1250	TGTGTGAGAGTGTGCTGATGACATCTGCTTGTTCACACATGACAGTTCTGTCTGG	1305		
Db 709	TGTGTGAGAGTGTGCTGATGACATCTGCTTGTTCACACATGACAGTTCTGTCTGG	650		
Qy 1310	GTCCTTATCAGGTGCCAAGCCCTGTCTCGGTCTGAGAGTACAGCAGTGAAGCAAGAG	1365		
Db 649	GTCCTTATCAGGTGCCAAGCCCTGTCTCGGTCTGAGAGTACAGCAGTGAAGCAAGAG	590		
Qy 1370	ACAAATATTCCTGCTCAGAGAGACAAACTGGCAATCCTTGAGCTGACACATCTTTC	1428		
Db 589	ACAAATATTCCTGCTCAGAGAGACAAACTGGCAATCCTTGAGCTGACACATCTTTC	530		
Qy 1430	ATGACCAATAGTCACTGTCTTACACTGGGTACATCTTGTACAGAGTGGCTCCACTGAT	1488		
Db 529	ATGACCAATAGTCACTGTCTTACACTGGGTACATCTTGTACAGAGTGGCTCCACTGAT	470		
Qy 1490	GCTGTGTCTCAGGCACTCTGTCCAAAGACAAATCCCTTTCACAAACCAAGCTGCTT	1544		
Db 469	GCTGTGTCTCAGGCACTCTGTCCAAAGACAAATCCCTTTCACAAACCAAGCTGCTT	410		
Qy 1550	TGTATCTTGTACCTTTTCAGAGAAAGGAGGTATCCCTGTGCAAAAGGCTCCAGGCTCT	1605		
Db 409	TGTATCTTGTACCTTTTCAGAGAAAGGAGGTATCCCTGTGCAAAAGGCTCCAGGCTCT	350		
Qy 1610	CCCTGCAACTCAGAGACCAAGCCACAGCTCACTCTGGAACTGTGTCCACAGATCTCTG	1665		
Db 349	CCCTGCAACTCAGAGACCAAGCCACAGCTCACTCTGGAACTGTGTCCACAGATCTCTG	290		
Qy 1670	TTCCTCTGATTAAGATTCCTCCTTCAAGGCTTAAGCTGAGATTTGGGCGACAGATAG	1728		
Db 289	TTCCTCTGATTAAGATTCCTCCTTCAAGGCTTAAGCTGAGATTTGGGCGACAGATAG	230		
Qy 1730	AATCAAACTATAGAGGTAGTCTTGTCTTAACTCAAGACTGTTCGTGAATAGGGTCCAG	1788		
Db 229	AATCAAACTATAGAGGTAGTCTTGTCTTAACTCAAGACTGTTCGTGAATAGGGTCCAG	170		
Qy 1790	GCTGTCTCAACATAGGGGCTTCTACCTGAGACCAAGGTTAGGAGACAGATTAAGCGAG	1848		
Db 169	GCTGTCTCAACATAGGGGCTTCTACCTGAGACCAAGGTTAGGAGACAGATTAAGCGAG	110		
Qy 1850	GTCGTGCTGTGGCACTGAGAAAGTCCACAGGTGGAGACTCTTCTGGGAGCACTTGGGGT	1909		
Db 109	GTCGTGCTGTGGCACTGAGAAAGTCCACAGGTGGAGACTCTTCTGGGAGCACTTGGGGT	51		
Qy 1910	CACAAATCCAGGTCATACCTTGAAGTTTGGATACCATGATATGATG 1958			
Db 50	BAAAATSCVGGTCATATGATNTAGTAT-T-GATTCGAGAGGAWAKRKR 3			

Homo sapiens cDNA clone CS0DL004YF04 3-PRIME, mRNA sequence.
 BX325344
 BX325344.1 GI:30336457
 EST.
 Homo sapiens (human)
 SOURCE
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 947)
 L1.W.B., Gruber,C., Jesssee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10000.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0AL004DC02NP1&cluster=10000.f. Contact :
 Peng Liang Email : filiang@lifetech.com URL :
 http://fulllength.invitrogen.com/ invitrogen Corporation 1600
 Paradey Avenue Genoscope sequence ID : CS0AL004DC02NP1.
 FEATURES
 SOURCE
 Location/Qualifiers
 1..947
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL004YF04"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_1lb="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /primer="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 232 a 241 c 241 g 205 t
 ORIGIN
 Query Match 43.3% Score 892; DB 13; Length 947;
 Best Local similarity 96.4% Pred. No. 1.4e-72;
 Matches 907; Conservative 22; Mismatches 10; Indels 2; Gaps 2;
 Oy 1030 CATCTCTGGAGGCTGCCAGGTGCTCCTGATGACTGGGCTGGACGAAGCCAGGAA 1089
 Db CTTCTCTGGAGGCTGCCAGGTGCTCCTGATGACTGGGCTGGACGAAGCCAGGAA 888
 Oy 1090 GAGAGGAGACAGAGAGAGTGTCTGATGAGTACGACCTTACTACACGCTCAGCT 1149
 Db RAGAGAGRAGAMAGAGAGTGTCTGATGAGTACGACCTTACTACACGCTCAGCT 828
 Oy 1150 GCCAAGAGTGAATGAGTGGGGGTAGCTCAAGCTTCTCTGCTCTGAAGTGGCC 1209
 Db GCCAAGAGTGAATGAGTGGGGGTAGCTCAAGCTTCTCTGCTCTGAAGTGGCC 768
 Oy 1210 AAACCTTCAGTCTCCAGAGGAGAAAGGACAGACCTGTGTGAGTGTGTGTGA 1269
 Db AAACCTTCAGTCTCCAGAGGAGAAAGGACAGACCTGTGTGAGTGTGTGTGA 708
 Oy 1270 CACATCTGCTTTGTTCCACACATGACGTTCTCTGCTGTGTGTGTGTGTGTGTGT 1329
 Db CACATCTGCTTTGTTCCACACATGACGTTCTCTGCTGTGTGTGTGTGTGTGTGT 648
 Oy 1330 CTTGTTCTCGGTCTGGAGATGACAGTGAAGCAATATTTCTCTGCTCAG 1389
 Db CTTGTTCTCGGTCTGGAGATGACAGTGAAGCAATATTTCTCTGCTCAG 588
 Oy 1390 AGATGACAACTGGCATCTTGAAGTGAACACTTTTCAATGACATAGTCACTGTCT 1449
 Db AGATGACAACTGGCATCTTGAAGTGAACACTTTTCAATGACATAGTCACTGTCT 528

Oy 1450 AACTGGGTACACTTTTGATCCAGATGTGGGCTTCCACTGATGTGTGTGTGTGTGTGT 1509
 Db AACTGGGTACACTTTTGATCCAGATGTGGGCTTCCACTGATGTGTGTGTGTGTGTGT 469
 Oy 1510 GTCCAGGACAAATCCCTTTCACAAACAAACAGCTGCTTTGTATCTTGTACCTTTTACG 1569
 Db GTCCAGGACAAATCCCTTTCACAAACAAACAGCTGCTTTGTATCTTGTACCTTTTACG 409
 Oy 1570 AGAAAGGAGGTATCCCTGTGTCGAAGGCTCCAGGCTCTCCCTGCAACTCAGAGCCCA 1629
 Db AGAAAGGAGGTATCCCTGTGTCGAAGGCTCCAGGCTCTCCCTGCAACTCAGAGCCCA 349
 Oy 1630 AGCCGAGTCACTCTGGGAACTGTGTCTCCAGCATCTGTCTCTTGAATTAAGATTC 1689
 Db AGCCGAGTCACTCTGGGAACTGTGTCTCCAGCATCTGTCTCTTGAATTAAGATTC 289
 Oy 1690 TCCTTCCAGGCTTAAGCTTGAGATTTGGGACAGATTAAGATCAATCTAAGAGCTAG 1749
 Db TCCTTCCAGGCTTAAGCTTGAGATTTGGGACAGATTAAGATCAATCTAAGAGCTAG 229
 Oy 1750 TTCTTGTCTAACTCAAGACTGTGTCTGAAATGAGGCTCCAGGCTTCAACCAATGGGCTT 1809
 Db TTCTTGTCTAACTCAAGACTGTGTCTGAAATGAGGCTCCAGGCTTCAACCAATGGGCTT 169
 Oy 1810 CTGACTGAGACCAAGGTTGAGGACAGATTAAGGACAGGCTGTCTCTGTGACCACTG 1869
 Db CTGACTGAGACCAAGGTTGAGGACAGATTAAGGACAGGCTGTCTCTGTGACCACTG 109
 Oy 1870 GAAATGCCAGATGAGGACTCTCTGGGAGCACTTGGGCTTCCAAATCCAGGTCCACT 1929
 Db GAAATGCCAGATGAGGACTCTCTGGGAGCACTTGGGCTTCCAAATCCAGGTCCACT 49
 Oy 1930 CTAGTTTGTGATAC-CATGATATGATGTATTAATCTGTGTC 1969
 Db CTAGTTTGTGATAC-CATGATATGATGTATGATTTTBKGTGS 8
 RESULT 6
 LOCUS 602597925P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4706575 5',
 DEFINITION mRNA sequence.
 ACCSSION B0575275
 VERSION B0575275.1 GI:13582928
 KEYWORDS
 ORGANISM
 Homo sapiens (human)
 SOURCE
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 975)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 CONTACT: Robert Strauberg, Ph.D.
 Email: gsa@bse-mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 cDNA Sequencing by: Incyte Genomics, Inc.
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lnl.gov
 Plate: LLM10575 row: a column: 08
 High quality sequence stop: 891.
 FEATURES
 SOURCE
 Location/Qualifiers
 1..975
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4706575"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_87"

/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 199 a 301 c 262 g 213 t
ORIGIN

Query Match 42.2%; Score 868.8; DB 10; Length 975;
Best Local Similarity 96.2%; Pred. No. 1.7e-70;
Matches 922; Conservative 0; Mismatches 32; Indels 4; Gaps 3;

713 GCCCAGCTTACGTGGAGACCTGAGATCTCAGCTCAGGAGCTCATCTCGGAGCAGGAG 772
1 GCCCAGCTTACGTGGAGACCTGAGATCTCAGCTCAGGAGCTCATCTCGGAGCAGGAG 60
773 CCGATGTCTGAGAGCTGGAGCGGTGCGCTCTTCAACATGTGAGCGGAGCCTACCCCTTCC 832
61 CCGATGTCTGAGAGCTGGAGCGGTGCGCTCTTCAACATGTGAGCGGAGCCTACCCCTTCC 120
833 AGGACTCGGAGCTGTCTCTGCTCTTTCGAGCAAGATCCGCGCGGAGCCTACCGCTTGCCTG 892
121 AGGACTCGGAGCTGTCTCTGCTCTTTCGAGCAAGATCCGCGCGGAGCCTACCGCTTGCCTG 180
893 CAGGCTCTCGGAGCTGTCTCTGCTCTTTCGAGCAAGATCCGCGCGGAGCCTACCGCTTGCCTG 952
181 CAGGCTCTCGGAGCTGTCTCTGCTCTTTCGAGCAAGATCCGCGCGGAGCCTACCGCTTGCCTG 240
953 AACGGCTCAGAGCCAGGAGCATCTCTGAGCAAGATCCGCGCGGAGCCTACCGCTTGCCTG 1012
241 AACGGCTCAGAGCCAGGAGCATCTCTGAGCAAGATCCGCGCGGAGCCTACCGCTTGCCTG 300
1013 TAGCCCAACCCGATCCCATCTCTGAGCAAGATCCGCGCGGAGCCTACCGCTTGCCTG 1072
301 TAGCCCAACCCGATCCCATCTCTGAGCAAGATCCGCGCGGAGCCTACCGCTTGCCTG 360
1073 TCGAGCAAGCCAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1132
361 TCGAGCAAGCCAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
1133 CTACTACAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 1192
421 CTACTACAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 480
1193 TGCCTCTGAACTGAGCCAAACCTTCACTGCTTCCAGAAAGGAGAAAGGAGAGCTGT 1252
481 TGCCTCTGAACTGAGCCAAACCTTCACTGCTTCCAGAAAGGAGAAAGGAGAGCTGT 540
1253 GTGAGGTGTGCTGTGTACACATCTGCTTGTTCACACACATGAGCTTCTGCTTGGGTG 1312
541 GTGAGGTGTGCTGTGTACACATCTGCTTGTTCACACACATGAGCTTCTGCTTGGGTG 600
1313 CTATACAGGTGCAAGCCCTGTCTCTGAGCTGAGGAGTACAGCAGTGAAGCAAGAGACA 1372
601 CTATACAGGTGCAAGCCCTGTCTCTGAGCTGAGGAGTACAGCAGTGAAGCAAGAGACA 660
1373 ATATTCCCTGCTCAGAGATGACAACTGGCATCTTGAAGTGAACAACCTTTCATG 1432
661 ATATTCCCTGCTCAGAGATGACAACTGGCATCTTGAAGTGAACAACCTTTCATG 720
1433 ACATAGGTCACTGTCTACACAGGAGTACCTT--GTAC--TGTGAGCTTCACTGATG 1490
721 ACATAGGTCACTGTCTACACAGGAGTACCTTGTGTACCAAGTGTGAGTCTTCACTGATG 780
1491 CTGAGTCTCAGGACCTTGTTCAGAGCAATCCCTTTCACAAACAAACAGCTTGCCTT 1550
781 CTGAGTCTCAGGACCTTGTTCAGAGCAATCCCTTTCACAAACAAACAGCTTGCCTT 840
1551 -GTATCTTGTACCTTTT--AGAGAAAGGAGTATCCCTGAGCAAGAGCTTCAAGCTTC 1608
841 GGTATCTGTACCTTTTCAAGATTAAGGAGAGAAATCCCTGAGGCAAGAGCTTCAAGCTTC 900
1609 TCCCTGCAACTCAGAGACCAAGCCAGCTCACTGTGGAACTGTGTTCCAGCATCT 1666

Db 901 TTCCCTGGAATCAAGACCCAGCCAGCTCACTCTGGAACCTGGGTTCCAGATCCT 958

RESULT 7

AL562576/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1001 bp mRNA linear EST 31-MAY-2003
AL562576 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC013J21 3-PRIME, mRNA sequence.
AL562576
AL562576.2 GI:31286588
EST.
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1001)
L1, M.B., Gruber, C., Jessup, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12911134.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10000.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC013J21&INPlicluster=10000.f. Contact :
Peng Liang Email: fliang@life.com/ invitrogen.com URL: Corporation 1600
Fareday Avenue Genoscope sequence ID : CS0DC013J21INP.
Location/Qualifiers
1..1001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC013J21"
/cbase_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

BASE COUNT 230 a 250 c 257 g 224 t
ORIGIN

Query Match 41.9%; Score 862.4; DB 9; Length 1001;

Best Local Similarity 94.0%; Pred. No. 6.5e-70;

Matches 942; Conservative 33; Mismatches 18; Indels 9; Gaps 8;

969 AGGATCTCTCTGACCCCTGCTGAGCAGAGACCCAGATGCTTACGCCCAACCCGATC 1028
1000 AAGCATCTCTCTGACCCCTGCTGAGCAGAGACCCATGCTTACGTTACCAACCCGATC 941
1029 CCATCTCTGAGAGCTGCTGAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1088
940 YATCTTGTGAGAGCTGCTGAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 881
1089 AGAGAGGAGAGACAGAGAGGAGTGTCTGATAGCTGAGACCAACCTTACTACAGCTGAC 1148
880 AGAGAGGAGAGACAGAGAGGAGTGTCTGATAGCTGAGACCAACCTTACTACAGCTGAC 821
1149 TGCCAACAGTGAATGATTTGGGGGTGAGCTCAAGCTTCTCTGCTGCTGAGCTGAGC 1208
820 TKCCAACAGTGAATGATTT--GGGGTGTGCTCAAGCTTCTCTGCTGCTGAGCTGAGC 762
1209 CAAACCTTCACTGCTTCCAGAGAGGAGAAAGCAAGAGCTGTGTGAGAGTGTGTGT 1268
761 CAAACCTTCACTGCTTCCAGAGAGGAGAAAGCAAGAGCTGTGTGAGAGTGTGTGT 702
1269 ACATATCTGCTTGTTCACACACACATGAGCTTCTGCTGAGCTTATCAGAGTGCACA 1327

QY 781 T-GAGACCTGAGCGCTGCTCTTCAACATGCTGCGCGGCACTACCCCTTCCAGAGCTC 839
 DB 783 TGGAGACCTGAGCGCTGCTCTTCAACATGCTGCGCGGCACTACCCCTTCCAGAGCTC 842
 QY 840 GAGACCTGCTGCTCTTCTTGGCAAGATCCGC 870
 DB 843 GGAGCCTGCTGCTCTTCTTGGAAAAAACCC 873
 RESULT 9
 LOCUS BQ941789 881 bp mRNA linear EST 21-APR-2002
 DEFINITION AGENCOURT 8821009 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6422656
 5', mRNA sequence.
 ACCESSION BQ941789
 VERSION BQ941789.1 GI:22357267
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 881)
 NIH-MGC http://mgs.nci.nih.gov/
 UNPUBLISHED
 CONTACT: Robert Strauberg, Ph.D.
 Email: rstraub@nihs.nih.gov
 Tissue Procurement: DCTD/DTF/Gazdar
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing By: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LICM2599 row: P column: 17
 High quality sequence from: 689.
 Location/Qualifiers
 1..881
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6422656"
 /csize_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_18"
 /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC library."
 BASE COUNT 161 a 275 c 262 g 181 t 2 others
 ORIGIN
 Query Match 41.1%; Score 846.2; DB 13; Length 881;
 Best Local Similarity 98.6%; Pred. No. 2.1e-68;
 Matches 863; Mismatches 0; Mismatches 10; Indels 2; Gaps 1;
 QY 531 GCTCTTCCGACATGAGCCAGCCCTTGGCGCATCTCAACAGCGTCTGCTGCG 590
 DB 1 GCTCTTCCGACATGAGCCAGCCCTTGGCGCATCTCAACAGCGTCTGCTGCG 60
 QY 591 TGATCTCAAGCTGTGTGCTTGTCTTCCGTGACCGTGAAGAGAAAGCTGTCTGA 650
 DB 61 TGATCTCAAGCTGTGTGCTTGTCTTGTGCTGACCGTGAAGAGAAAGCTGTCTGA 120
 QY 651 GAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 710
 DB 121 GAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

QY 711 GTGCCAGCCTTACGTGGACCTGAGATCTCAGCTCAGCGGCTCATACTCGGGCAAGGC 770
 DB 181 GTGCCAGCCTTACGTGGACCTGAGATCTCAGCTCAGCGGCTCATACTCGGGCAAGGC 240
 QY 771 AGCCGATGTGAGAGCCTGGGCGTGGAGCTCTTCAACATGCTGCGCGGCACTACCCCTT 830
 DB 241 AGCCGATGTGAGAGCCTGGGCGTGGAGCTCTTCAACATGCTGCGCGGCACTACCCCTT 300
 QY 831 CCAGGACTCGAGAGCCTGTCTCTTCTTGGCAAGATCCGCGCGGCGGCTTACGCTTGC 890
 DB 301 CCAGGACTCGAGAGCCTGTCTCTTCTTGGCAAGATCCGCGCGGCGGCTTACGCTTGC 360
 QY 891 TGCAGGCTCTGCGGCGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950
 DB 361 TGCAGGCTCTGCGGCGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 951 TGAACGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1010
 DB 421 TGAACGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 480
 QY 1011 CTTAGCCCAACCCGATCCCATCTCTGAGAGCTGCGCCAGGTGATCCCTGATGACTGCG 1070
 DB 481 CTTAGCTCAACCCGATCCCATCTCTGAGAGCTGCGCCAGGTGATCCCTGATGACTGCG 540
 QY 1071 GCTGACGAGAGCAGGAG 1130
 DB 541 GCTGACGAGAGCAGGAG 600
 QY 1131 CCTGACTCAAGCTGAGCTGAGCAAGTGAATGATTTGAGGAGAGTCCAGAGCTTCT 1190
 DB 601 CCTGACTCAAGCTGAGCTGAGCAAGTGAATGATTTGAGGAGAGTCCAGAGCTTCT 660
 QY 1191 CCTGCTCTGAACTGAGCAAACTTCACTGCTTCCAGAGAGAGAGAGAGAGAGAGCT 1250
 DB 661 CCTGCTCTGAACTGAGCAAACTTCACTGCTTCCAGAGAGAGAGAGAGAGAGAGCT 720
 QY 1251 GTGTGAGATGT 1310
 DB 721 GTGTGAGATGT 780
 QY 1311 TGCTTATCAGGAGGAGCCCTGTTCTCGGTGCTGAGAGTACAGAGTGAAGAGAGAGA 1370
 DB 781 TGCTTATCAGGAGGAGCCCTGTTCTCGGTGCTGAGAGTACAGAGTGAAGAGAGAGA 840
 QY 1371 CAATATTCCTCTGCTCAAGAGATGACAACTGG 1403
 DB 841 CAATATTCCTCTGCTCAAGAGATGACAACTGG 875
 RESULT 10
 BX325345 945 bp mRNA linear EST 02-MAY-2003
 LOCUS BX325345 Homo sapiens B CELLS (RAMOS CELL LINE) COR 25-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS00D0041P04 5-PRIME, mRNA sequence.
 ACCESSION BX325345
 VERSION BX325345.1 GI:30338449
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 945)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 UNPUBLISHED
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: beauref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10000.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/

FEATURES	source
ORGANISM	Homo sapiens (human)
REFERENCE	1. (bases 1 to 863)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgarbde-rc@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LNCMI676 row: 1 column: 13 High quality sequence atop: 833. Location/Qualifiers 1. 863 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4842324" /tissue_type="normal pigmented retinal epithelium" /lab_host="DH10B (phage-resistant)"

Query Match	Similarity	Score	DB 10	Length	863
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 863	
Best Local Match	98.2%	Pred. No. 6.6e-66			
Matches	838	Conservative	0	Mismatches	14
				Indels	1
				Gaps	1
BASE COUNT	135 a	300 c	264 g	164 t	
ORIGIN					
Query Match	39.8%	Score 818.6	DB 10	Length 8	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:09:04 ; Search time 583.591 Seconds

(without alignment)

9524.031 Million cell updates/sec

Title: US-09-909-474D-1

Perfect score: 2059

Sequence: 1 gctctggcccgccgagcgc.....aaaaaaaaaaaaaaaaa 2059

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /SIDSL/gcgdata/geneseq/geneeqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA2001.DAT.*
23: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneeq/geneeqn-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Hit	Score	Query Match	Length	DB ID	Description
1	1999.2	97.1	2116	22	AAAF30480
2	1997.2	97.0	2059	22	AAAF59850
3	1997.2	97.0	2092	22	AAAF58064
4	1767.8	85.9	2048	21	AAAF7866
5	1073.8	52.2	1077	22	AAAF6709
6	1054.6	51.2	1083	24	ABN86479
7	1004.4	48.8	1085	21	AAAF6155
8	802.2	39.0	972	22	AAAF6218

9	645.2	31.3	1076	24	ABL39747	Human NS CDNA sequ
10	645.2	31.3	1076	24	ABL39762	Human NS CDNA sequ
C 11	562.4	27.3	996	23	AAS91231	DNA encoding novel
12	533	25.9	541	24	ABK5110	Human colon cancer
C 13	505.2	24.5	2562	16	AAO89817	DNA encoding the A
C 14	426	20.7	426	24	ABL67304	Human ovarian can
15	330.2	16.0	396	22	AAS37461	Novel human diagn
C 16	315.4	15.3	353	24	ABL62004	Colon adenocarcino
17	296	14.4	534	23	AAS68600	DNA encoding novel
18	274.2	13.3	3319	24	ABN86478	Human tribbles hom
19	268.8	13.1	562	21	AAAF8131	Human cancer assoc
20	263.2	12.8	1909	22	AAS45035	CDNA encoding nove
21	251.4	12.2	3280	25	ABK32619	Human CDNA #269 di
22	247	12.0	3324	24	AAS94913	Human DNA sequenc
23	242.6	11.8	690	24	ABT09024	Phase-1 Rat CT gen
24	234	11.4	675	24	ABK63550	Human CDNA differe
25	233.4	11.3	942	25	ABX74428	Human CDNA sequenc
26	232.4	11.3	818	25	ACG45125	Human C8FW encodin
27	222.2	10.8	621	25	ACG45126	Human C8FW encodin
28	207.6	10.1	1943	22	AAS45223	CDNA encoding nove
C 29	205	10.0	207	21	AAAF3717	Human secreted pro
C 30	165	8.0	498	24	ABK63084	Human immune/haema
C 31	163.6	7.9	7282	22	AAAF6318	Human immune/haema
C 32	158	7.7	562	22	AAAF1922	Human brain expres
C 33	158	7.7	562	22	AAAF10231	Human bone marrow
C 34	158	7.7	562	22	AAAF6130	Human bone marrow
C 35	158	7.7	562	22	AAAF1844	Human liver single
C 36	158	7.7	562	23	ABK55823	Human genome-deriv
C 37	158	7.7	562	24	ABK10267	Human immune/haema
C 38	158	7.7	25772	22	AAAF1332	Human foetal liver
C 39	156	7.6	417	22	ABAF4417	Human foetal liver
C 40	156	7.6	417	22	AAAF2879	Human brain expres
C 41	156	7.6	417	22	AAAF4952	Human bone marrow
C 42	156	7.6	417	22	AAAF4879	Human bone marrow
C 43	156	7.6	417	23	ABK48709	Probe #21565 used
C 44	156	7.6	417	24	ABK48709	Human liver single
C 45	142.8	6.9	2038	23	ABL29125	Human genome-deriv
						Drosophila melanog

ALIGNMENTS

```
RESULT 1
ID AAF30480 standard; cDNA; 2116 BP.
AC AAF30480;
XX
DT 29-MAY-2001 (first entry)
XX
DB Human protein phosphatase and kinase protein-5 CDNA 1271505CB1.
XX
XX Protein phosphatase and kinase protein; PPHK-P-5; human;
KW gastrointestinal disorder; immune system disorder;
KW neurological disorder; cell proliferative disorder; cancer;
KW diagnostic therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key 161.1237 Location/Qualifiers
XX FT CDS /*caga a
XX FT misc_feature 326..385 /*caga b
XX FT /*note= "unique fragment"
XX
XX W0200120004-A2.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000MO-US25515.
XX CDNA J503-KS enco
XX
XX 15-SEP-1999; 99US-0154141.
```

XX (INCY-) INCYTE GENOMICS INC.
 PA
 PI Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzal Y,
 PI Lu DM,
 XX MPI, 2001-244811/25.
 DR P-PSDB; AAB20326.
 XX
 PT Novel human protein phosphatase and kinase proteins for diagnosis,
 PT treatment and prevention of gastrointestinal, immune system,
 PT neurological and cell proliferative disorders -
 XX
 PS Claim 5; Page 99; 103pp; English.
 XX
 CC The present sequence is that of cDNA encoding novel human
 CC protein phosphatase and kinase protein PPHK-5 (see AAB20326).
 CC The cDNA was initially identified in Incyte Clone ID No. 1271505CB1,
 CC from a foetal lung tissue library. The PPHK-5 gene was localised
 CC to human chromosome 20. Tissues that express PPHK-5 (as a
 CC fraction of total tissues expressing PPHK-5) include reproductive
 CC (0.288), gastrointestinal (0.212) and haematopoietic or immune
 CC (0.192). Diseases or conditions associated with tissues expressing
 CC PPHK-5 (as a fraction of total tissues expressing PPHK-5) include
 CC cancer (0.577), inflammation or trauma (0.327) and cell
 CC proliferation (0.308). The encoded protein shows homology to rat
 CC kinase. The invention provides human PPHK-1 to -11 polypeptides
 CC (see AAB20322-32) and polynucleotides (see AAB20476-86). It also
 CC provides expression vectors, host cells, antibodies, agonists
 CC and antagonists, as well as methods for diagnosing, treating or
 CC preventing disorders associated with expression of PPHK, including
 CC gastrointestinal disorders, immune system disorders, neurological
 CC disorders and cell proliferative disorders, including cancer.
 CC
 SQ Sequence 2116 BP; 407 A; 663 C; 609 G; 437 T; 0 other;
 Query Match 97.1%; Score 1999.2; DB 22; Length 2116;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2001; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DB 593 CACAGCTGTGGAGAACGCGCACCGATCCCTGAGCTGAGGCTGCGCTCTTCCGC 652
 QY 541 CAGATGGCCACCGCCCTGGCGCACTGTCAACAGACGGTGTGCTCTGATCTCAAG 600
 DB 653 CAGATGGCCACCGCCCTGGCGCACTGTCAACAGACGGTGTGCTCTGATCTCAAG 712
 QY 601 CTGTGTGCTTGTCTTGTGCTGACCGTGAAGAGAAAGCTGTGTGAAGAACTTGAG 660
 DB 713 CTGTGTGCTTGTCTTGTGCTGACCGTGAAGAGAAAGCTGTGTGAAGAACTTGAG 772
 QY 661 GACTCTGCGTGTGCTGATGGGCGCAGATATCCCTGTGGGCAAGCAACGGTGGCCAGCC 720
 DB 773 GACTCTGCGTGTGCTGATGGGCGCAGATATCCCTGTGGGCAAGCAACGGTGGCCAGCC 832
 QY 721 TAGCTGGAGCTGAGATCACTCACTCAAGGAGCTCATATCTGAGGCAAGGACGAGATGTC 780
 DB 833 TAGCTGGAGCTGAGATCACTCACTCAAGGAGCTCATATCTGAGGCAAGGACGAGATGTC 892
 QY 781 TGAAGCTGGAGCTGTGGCTCTTCAACATGTGGCGGCACTAACCCCTTCCAGAACTCG 840
 DB 893 TGAAGCTGGAGCTGTGGCTCTTCAACATGTGGCGGCACTAACCCCTTCCAGAACTCG 952
 QY 841 GAGCTGTCTGTCTTGTGCGCAAGATCCGCGCGGGGCTTAACGCTTGGCTGACGGCTC 900
 DB 953 GAGCTGTCTGTCTTGTGCGCAAGATCCGCGCGGGGCTTAACGCTTGGCTGACGGCTC 1012
 QY 901 TCGGCCCCCTGCGCTGTCTGTGCTGTGCTCTCTTGTGTGGAGCAAGCTGAACGCTC 960
 DB 1013 TCGGCCCCCTGCGCTGTCTGTGCTGTGCTCTCTTGTGTGGAGCAAGCTGAACGCTC 1072
 QY 961 ACAGCCACAGGCAATCTCTCTGACACCTGTGCTGACAGAACCCGATGCTTACGCCCA 1020
 DB 1073 ACAGCCACAGGCAATCTCTCTGACACCTGTGCTGACAGAACCCGATGCTTACGCCCA 1132
 QY 1021 ACCGATCCGATCTGTGGAGGCGCTGCCAGGTGTCTCATATGATCGGGCTGAGAGAA 1080
 DB 1133 ACCGATCCGATCTGTGGAGGCGCTGCCAGGTGTCTCATATGATCGGGCTGAGAGAA 1192
 QY 1081 GCCAGGAGAGAGAGAGAGACAGAGAAAGTGTCTGTATGATAGACCAACCTTACTACA 1140
 DB 1193 GCCAGGAGAGAGAGAGAGACAGAGAAAGTGTCTGTATGATAGACCAACCTTACTACA 1252
 QY 1141 CGCTCAAGCTCCCAACATGTGATTTGATTTGGGGGTAGCTCCAGCCTTCTCTGCTCTG 1200
 DB 1253 CGCTCAAGCTCCCAACATGTGATTTGATTTGGGGGTAGCTCCAGCCTTCTCTGCTCTG 1312
 QY 1201 AACTGAGCCAAACCTTCAAGTCCCTCCAGAAAGGAGAAAGCCAGTGTGGAGTG 1260
 DB 1313 AACTGAGCCAAACCTTCAAGTCCCTCCAGAAAGGAGAAAGCCAGTGTGGAGTG 1372
 QY 1261 TGTGTGTATACATCTGTCTTGTTCACACATGACAGTTCCTGTGGGTCTTATCAG 1320
 DB 1373 TGTGTGTATACATCTGTCTTGTTCACACATGACAGTTCCTGTGGGTCTTATCAG 1432
 QY 1321 GTGCAAGCCCTGTCTTGTGCTGTGGAGTACAGACGTGACAAAGAGACAAATATTCCTC 1380
 DB 1433 GTGCAAGCCCTGTCTTGTGCTGTGGAGTACAGACGTGACAAAGAGACAAATATTCCTC 1492
 QY 1381 TGTCTCAAGAGATGACAAACTGGATCTTGAAGCTGACAACTTTTCAATGACATAGG 1440
 DB 1493 TGTCTCAAGAGATGACAAACTGGATCTTGAAGCTGACAACTTTTCAATGACATAGG 1552
 QY 1441 TCACTGTCTACACTGTGGTACACTTTGTATCAAGTGTGGCTCTCACTGATCTGTGTCTCA 1500
 DB 1553 TCACTGTCTACACTGTGGTACACTTTGTATCAAGTGTGGCTCTCACTGATCTGTGTCTCA 1612
 QY 1501 GGCACCTGTCTCAAGAGAACTCCCTTTACAAACCAACGAGTGTCTTGTATCTTGTAT 1560
 DB 1613 GGCACCTGTCTCAAGAGAACTCCCTTTACAAACCAACGAGTGTCTTGTATCTTGTAT 1672
 QY 1561 CCTTTTCAAGAGAAAGGATATCCCTGTGCAAAAGGCTCAAGGCTCTCCCTGCAACT 1620

Db 1673 CCTTTGAGAGAAAGGAGATATCCCTGTGCAAGGCTCCAGGCTCTCCCTGCAACT 1732
 Qy 1621 CAGAACCCAGAGCCAGCTCACTCTGGGAACTGTGTTCCCAAGATCTGTCTCTTGAAT 1680
 Db 1733 CAGAACCCAGAGCCAGCTCACTCTGGGAACTGTGTTCCCAAGATCTGTCTCTTGAAT 1792
 Qy 1681 AAGAGATTCCTCTCCAGGCTCAAGGCTGGGATTTGGGCAAGATTAAGATCCAAACTA 1740
 Db 1793 AAGAGATTCCTCTCCAGGCTCAAGGCTGGGATTTGGGCAAGATTAAGATCCAAACTA 1852
 Qy 1741 TGAGGCTAGTCTTGTCTTAATCAAGACTGTTCTGGAATGAGGGTCCAGGCTGTCAAC 1800
 Db 1853 TGAGGCTAGTCTTGTCTTAATCAAGACTGTTCTGGAATGAGGGTCCAGGCTGTCAAC 1912
 Qy 1801 ATGGGGCTTCTGACCTGACGACCAAGGTTGAGGAGACAGATTAAGGACAGGCTGTCTGT 1860
 Db 1913 ATGGGGCTTCTGACCTGACGACCAAGGTTGAGGAGACAGATTAAGGACAGGCTGTCTGT 1972
 Qy 1861 GGCCACCTGAGAAAGTCCAGAGGTGGGACTCTTCTGGGAGACACTGGGGGTCCAATCCAG 1920
 Db 1973 GGCCACCTGAGAAAGTCCAGAGGTGGGACTCTTCTGGGAGACACTGGGGGTCCAATCCAG 2032
 Qy 1921 GTCCATACCTGAGGTTTGAATACCATGATGATGATTTACCTGTGCTTAATTAAGA 1980
 Db 2033 GTCCATACCTGAGGTTTGAATACCATGATGATGATTTACCTGTGCTTAATTAAGA 2092
 Qy 1981 GAATTAATGAATTAATAAAAAAAAA 2004
 Db 2093 GAATTAATGAATTAATAAAAAAAAA 2116

RESULT 2

AI59850
 ID AI59850 standard; cDNA; 2059 BP.

AAI59850/

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 3839.

Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 leukemia; B6.

Homo sapiens.

MO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US34263.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0552317.

09-JUL-2000; 2000US-0598042.

19-JUL-2000; 2000US-0620312.

03-AUG-2000; 2000US-0653450.

14-SEP-2000; 2000US-0662191.

19-OCT-2000; 2000US-0683036.

29-NOV-2000; 2000US-0727344.

(HYSB-) HYSBQ INC.

Tang YT, Liu C, Aundi V, Chen R, Ma Y, Qian XB, Ren P, Wang D,

Wang JT, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J,

Zhao QA, Zhou P, Goodrich R, Drmanac RT,

WPI, 2001-442253/47.

P-PBDB; AAM40694.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 3839; 10078bp; English.
 CC The invention relates to human nucleic acids (AA15798-AA161369) and
 CC the encoded polypeptides (AAM38642-AA42213) with nocotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activity, chemotactic/chemokinetic activity, haemostatic
 CC Activin/inhibin activity, cancer diagnosis and therapy, drug screening,
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SQ Sequence 2059 BP; 400 A; 645 C; 583 G; 431 T; 0 other;
 XX
 Query Match 97.0%; Score 1997.2; DB 22; Length 2059;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1999; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 GCTCTGAGCCCGGCGGGGCGCCGAGCGGAAAGAGCGGGGAGATGCGAGCC 60
 Db 58 GCTCTGAGCCCGGCGGGGCGCCGAGCGGAAAGAGCGGGGAGATGCGAGCC 117
 Qy 61 CCTCTGAGCTCTCTGCGGGGTTCCCTGTCCAGAAAGACGGTTGAGTTGATGACAC 120
 Db 118 CCTCTGAGCTCTCTGCGGGGTTCCCTGTCCAGAAAGACGGTTGAGTTGATGACAC 177
 Qy 121 TTAGATACCGAGGCTCCCGTCCAGAAACGAGCTGAAAGTGGGCCCCAGCACTGCC 180
 Db 178 TTAGATACCGAGGCTCCCGTCCAGAAACGAGCTGAAAGTGGGCCCCAGCACTGCC 237
 Qy 181 CCCGCGCTGTGAGCCCTGAGGCGCACTAGCTCCAGATGCTGGAACGCTGTGGCCACT 240
 Db 238 CCCGCGCTGTGAGCCCTGAGGCGCACTAGCTCCAGATGCTGGAACGCTGTGGCCACT 297
 Qy 241 GCTCTCCGCTTGGGCGCTTATGTCCTCTGAGGCCCGAGAGGGCGGGCGCTTACCG 300
 Db 298 GCTCTCCGCTTGGGCGCTTATGTCCTCTGAGGCCCGAGAGGGCGGGCGCTTACCG 357
 Qy 301 GCCCTGACCTGCTTACAGGCACTGATTAATCTGCAAGGTGTACCCCGTCAAGAGCC 360
 Db 358 GCCCTGACCTGCTTACAGGCACTGATTAATCTGCAAGGTGTACCCCGTCAAGAGCC 417
 Qy 361 CTGGCGGTGAGAGCCCTTACGCGGGGCTGGCCCGGCAAGAGATGTGGCTCGGCCACT 420
 Db 418 CTGGCGGTGAGAGCCCTTACGCGGGGCTGGCCCGGCAAGAGATGTGGCTCGGCCACT 477
 Qy 421 GAGGTCTGCTGATCCAGCTCTCTTACGCGCTTTTCTCTGCAAGCCATGGAACATG 480
 Db 478 GAGGTCTGCTGATCCAGCTCTCTTACGCGCTTTTCTCTGCAAGCCATGGAACATG 537
 Qy 481 CACAGCTGTGAGAGCCCGCACCGTATCCCTGAGGCTGAGGCTGGCGTCTTCCGC 540
 Db 538 CACAGCTGTGAGAGCCCGCACCGTATCCCTGAGGCTGAGGCTGGCGTCTTCCGC 597
 Qy 541 CAGATGCGACGCGCTGCGGCACTGTACACAGACGATGATGATGATGATGATGATG 600
 Db 598 CAGATGCGACGCGCTGCGGCACTGTACACAGACGATGATGATGATGATGATGATG 657
 Qy 601 CTGTGTGCTTTGTCTTGTGCTGACCGTGAAGAGAAAGCTGTGTCTGAAGAACTTGG 660
 Db 658 CTGTGTGCTTTGTCTTGTGCTGACCGTGAAGAGAAAGCTGTGTCTGAAGAACTTGG 717

OY	661	TACTCTGCGTGCATGACCTGGGCGCAATGATTTCCCTGTGGGAAACAACA	CGCGTGGCCAGCC	720
Db	718	GACTCTGTGTGTCTGATCTGTGGCGCAATGATTTCCCTGTGGGAAACAACA	CGCGTGGCCAGCC	777
OY	721	TACGTGGGACCTGAGATATCTCAGCTCA	CGGAGCTATACCTGGGCAAGCAGCCGATGTC	780
Db	778	TACGTGGGACCTGAGATATCTCAGCTCA	CGGAGCTATACCTGGGCAAGCAGCCGATGTC	837
OY	781	TGGAGCCTGGGGGTGGCCGCTCTTTCACATGCTGTGGCGCGCACTTACCCCTTCCAGGACCTGCG	840	
Db	838	TGGAGCCTGGGGGTGGCCGCTCTTTCACATGCTGTGGCGCGCACTTACCCCTTCCAGGACCTGCG	897	
OY	841	GAGCCTGTCTGTCTCTTCTGGCAAGAATCCGCGCGGGGGGCTACAGCCTTGCTGCAAGCCTC	900	
Db	898	GAGCCTGTCTGTCTCTTCTGGCAAGAATCCGCGCGGGGGGCTACAGCCTTGCTGCAAGCCTC	957	
OY	901	TCGGCCCCCTGGCCGCTGTCTGTGGTTGCGTGTGCTCTTTCGTCCGGAGCCAGCTGAACGCGCTC	960	
Db	958	TCGGCCCCCTGGCCGCTGTCTGTGGTTGCGTGTGCTCTTTCGTCCGGAGCCAGCTGAACGCGCTC	1017	
OY	961	ACAGGCAACAGGCAATCCTCTCTGCAACCCCTGGCGTGGGACAGAAACCGGATGCCCTTACGCCCA	1022	
Db	1018	ACAGGCAACAGGCAATCCTCTCTGCAACCCCTGGCGTGGGACAGAAACCGGATGCCCTTACGCCCA	1077	
OY	1021	ACCCGATCCCATCTCTGTGGAGGCGTGGCCCAAGTGTGCTCCCTGATGGA	CTGTGGGCTGTGACGAA	1086
Db	1078	ACCCGATCCCATCTCTGTGGAGGCGTGGCCCAAGTGTGCTCCCTGATGGA	CTGTGGGCTGTGACGAA	1137
OY	1081	GCCAGGGGAAAGAGAGGGAGACAGAGAAATGTGTTCTGTATGTGTATAGAACACACCTTACTTCA	1144	
Db	1138	GCCAGGGGAAAGAGAGGGAGACAGAGAAATGTGTTCTGTATGTGTATAGAACACACCTTACTTCA	1197	
OY	1141	CGCTCAGCTGCGCAACAGTGGATTTGAGTTTGGGGGTATAGCTCCAAAGCCTTCTCCTGTGCTGTG	1200	
Db	1198	CGCTCAGCTGCGCAACAGTGGATTTGAGTTTGGGGGTATAGCTCCAAAGCCTTCTCCTGTGCTGTG	1257	
OY	1201	AACTGAGCCMAACCTTCAATGTGCTTCCAGAAAGGAGAAAGGCATGCTGTGTGTGATGTG	1266	
Db	1258	AACTGAGCCMAACCTTCAATGTGCTTCCAGAAAGGAGAAAGGCATGCTGTGTGTGATGTG	1317	
OY	1261	TGCTGTGTACACATCTGTGCTTGTGTTCCACAACATGCAAGTCTGTGCTTGGGTGCTTATGAG	1320	
Db	1318	TGCTGTGTACACATCTGTGCTTGTGTTCCACAACATGCAAGTCTGTGCTTGGGTGCTTATGAG	1377	
OY	1321	GTGCCAGCCCTGTGTTCTGTGTGTCTGGGTGCTGGAGTACAGCATGTGTGACMAAAGAACATATTTCC	1380	
Db	1378	GTGCCAGCCCTGTGTTCTGTGTGTCTGGGTGCTGGAGTACAGCATGTGTGACMAAAGAACATATTTCC	1437	
OY	1381	TGCTTCACAGAGATGACAAATCTGGGACCTGTGAGCTGACAAACCTTTTCATGACCATAGAG	1446	
Db	1438	TGCTTCACAGAGATGACAAATCTGGGACCTGTGAGCTGACAAACCTTTTCATGACCATAGAG	1497	
OY	1441	TCACTGTGTACATCTGGGTATCACTTTGTATCAAGTGTCCGCTTCCACTGATGTGCTGTGCTCA	1500	
Db	1498	TCACTGTGTACATCTGGGTATCACTTTGTATCAAGTGTCCGCTTCCACTGATGTGCTGTGCTCA	1557	
OY	1501	GGCACCTGTGTTCMAAGAACATCCCTTTTCACAAACAAACCAAGCTGCTTTGTATTTGTGTA	1566	
Db	1558	GGCACCTGTGTTCMAAGAACATCCCTTTTCACAAACAAACCAAGCTGCTTTGTATTTGTGTA	1617	
OY	1561	CCTTTTTCAGAGAAAGGAGAGTATCCCTGTGCAAAAGGCTCCAGGCGCTCTCCGCTGCAACT	1628	
Db	1618	CCTTTTTCAGAGAAAGGAGAGTATCCCTGTGCAAAAGGCTCCAGGCGCTCTCCGCTGCAACT	1677	
OY	1621	CAGGACCCAGAGCCAGCTCACTCTGTGGAACTGTGTTCACAGCATTTCTGTCTCTTGATTT	1680	
Db	1678	CAGGACCCAGAGCCAGCTCACTCTGTGGAACTGTGTTCACAGCATTTCTGTCTCTTGATTT	1737	
OY	1681	AAAGAGATTTCTCTTTCAGAGGCTTAAGCCTGTGGGATTTTGGGCAAGATTAAGATTCMAACTA	1746	
Db	1738	AAAGAGATTTCTCTTTCAGAGGCTTAAGCCTGTGGGATTTTGGGCAAGATTAAGATTCMAACTA	1797	
OY	1741	TGAGGCTAGTTCTTGTCTTAACTCAAGACTGTGTTCTGGAAATGAGAGGTCCAGGCGCTGTCAAC	1800	

Db	1798	TGAGGCTAGTCTTGCTCAACTGATCTTGGAAATAGAGGATCCAGGCTGTCAACC	1857
Qy	1801	ATGAGGCTTGTGACCTGAGCACAAGGTTAGGACAGGATTAGGACAGGATCTGCTCTGT	1860
Db	1858	ATGGGGCTTCTGACCTGAGCACAAGGTTAGGACAGGATTAGGACAGGATCTGCTCTGT	1917
Qy	1861	GGCCACCTGGAAAGTCCAGGTGGGACCTCTCTGAGGACACTTGGGGATCCACATCCGAG	1920
Db	1918	GGCCACCTGGAAAGTCCAGGTGGGACCTCTCTGAGGACACTTGGGGATCCACATCCGAG	1977
Qy	1921	GTCCACTACTTAGGTTTGGATATCAGATGATGATATGTTTACCTGTGCTTAATAAGGA	1980
Db	1978	GTCCACTACTTAGGTTTGGATATCAGATGATGATATGTTTACCTGTGCTTAATAAGGA	2037
Qy	1981	GAATTATGAAATTAATAAAAAA 2002	
Db	2038	GAATTATGAAATTAATAAAAAA 2059	
RESULT 3			
AA158064	ID	AA158064 standard; cDNA; 2092 BP.	
XX	AA158064;		
AC			
XX			
DT	22-OCT-2001	(first entry)	
XX			
DE	Human polynucleotide SEQ ID NO 267.		
XX			
KW	Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemocytic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukemia; se.		
XX			
OS	Homo sapiens.		
XX			
PN	WO20015312-A1.		
XX			
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US34263.		
XX			
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0588042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSB-) HYSBQ INC.		
XX			
P1	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
P1	Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
P1	Zhao QH, Zhou P, Goodrich R, Drmanec RT;		
XX			
XX	WPI; 2001-442253/47.		
DR	P-P8DB; AAM38908.		
XX			
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX			
PS	Claim 1, SEQ ID NO 267, 10076bp; English.		
XX			
CC	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	the encoded polypeptides (AAM38642-AAM42213) with nocotropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		

system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Notes: The sequence data for this patent did not form part of the printed specification.

Sequence 2092 BP; 405 A; 658 C; 593 G; 436 T; 0 other;

Query Match	97.0%	Score 1997.2	DB 22	Length 2092
-------------	-------	--------------	-------	-------------

Matches 1999, Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCTCTGAGACCCCGGCGCGCCCGGCGCCACGCGGAAAGAGGGGAGATGCGAGCCACC 60
Db 91 GCTCTGAGACCCCGGCGCGCCCGGCGCCACGCGGAAAGAGGGGAGATGCGAGCCACC 150

6. CCTCTGGCTCTCTGCGGGTTCCTGTCGAGGAAGCGATTGAGTGGATGCAAC 120
 151 CCTCTGGCTCTCTGCGGGTTCCTGTCGAGGAAGCGATTGAGTGGATGCAAC 210

Qy 121 TTAGTATCCGAGCGTCCCGTCCAGAAAAGAGCTCGAAGTGGCCCGAGCCAGACTGGCCC 180
 Db 211 TTAGATACCGAGCGTCCCGTCCAGAAAAGAGCTCGAAGTGGCCCGAGCCAGACTGGCCC 270

07 CCGTGGCTGTGCCCCGTGAGCCCACTACTGCTCAGATGTGCAATGCTGTGGCACT 240
 18 CCGTGGCTGTGCCCCGTGAGCCCACTACTGCTCAGATGTGCAATGCTGTGGCACT 240
 Db 271 CCGTGGCTGTGCCCCGTGAGCCCACTACTGCTCAGATGTGCAATGCTGTGGCACT 330

331 GCGTCCCGTCTGGGCGCCCTATGTCCTCCCTGAGCCGAGAGGGCGGGCGGGCTACCAAG 390

Vy G C C C T G C A C T G C C C T A C A G G A C T A G A T A T C C T G A A G G T G T A C C C C G C C A G A A G C 350
 Db G C C C T G C A C T G C C C C T A C A G G A C T G A G T A T A C T G C A G G G T G T A C C C C G C C A G A A G C 450

DB 451 CTGGCGGTGTGAGCCCTATGCGGGCTGCCCCGACACAGCATGTGGCTGGCCACT 510

511 GAGGTCTGGCTGTACCCAGCTCCTCTACGCTTTTTCACTCGGACCCCATGGGACATG 570

571 CACAGCCTGTCGAGAGCCGCCAACCGTATCCTGAGCCTGAGGCTGCGGCTCTTCCGC 630

631 CAGATGGCCACCGCCCTGGGGCACTGTCAACAGACGGTGTGTTCTCTGCTGATCTCAAG 690

691 CTGTGTCCGTTTGTCTTCGTCGACCGTAGAGAGAAAGCTGTGTCTGGAGAACCTGGAG 750

Db 751 GACTCCTGCTGACTGCGGGCCAGATGATTCCCTGTGGGACAAGCACCAGTCCAGCC 810

Db	811	TACGTGGGACCTGAGATACTCTCAGCTCAACGGCCCTCATACTCGGGCAGAGCCGATGTC	870
Ov	781	TGGAGGCTCGGGGCGTGGGCGCTTTCACCATGTGTGGCGGGCCATPACCCCTTTCAGGACTGG	840

Db 871 TGGAGCCTGGGCGTGGCCCTCTTCAACCACTGGCCGGCCACTACCCCTTCCAGGACTCG 930

Qy 841 GAGCTGTCTCTGCTCTTTCGGGAAGATCGCCGCGGGGCTACAGCGCTTGGCTGCAAGGCTC 900

Db 931 GAGCTGTCTCTGCTCTTTCGGGAAGATCGCCGCGGGGCTACAGCGCTTGGCTGCAAGGCTC 990

Qy 901 TCGGCCCCCTGACCGCTGTCTGGTTCGCTGCTCTTGGTGGTGGAGACGAGCTGAACGGCTC 960

Db 991 TCGGCCCCCTGACCGCTGTCTGGTTCGCTGCTCTTGGTGGTGGAGACGAGCTGAACGGCTC 1050

Qy 961 ACGGCGACGAGGATCTCTGACCCCTGCTGCTGCAAGACGACCGGATGCGCTTAAGCCCA 1020

0222 ACCGATCCATCTCTGGAGGCTGCCAGGTGTCTCTGATGACTGGGGCTGGACGAA 1080

1081 GCCACGGGAAGAGGAGGAGACAGAGAAAGTGGTTCTGTATGGCTTAGGACACACCTACTACA 1140

1141 CGCTCAGCTGCCAAGTCGATTCGATTGGGGTAGCTCCAAAGCTTCTCCGCTCTCG 1200

1201 AACTGAGCCAAACCTTCAGTGCCTTCAGAGGAGAAAGGACGAGGCTGTGGAATG 1260
 1202 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
 1203 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||

[illegible]

Oy 1321 GTGCCAAGCCCTGTTCTCGTGCTGGAGTACAGCAGTGAAGCAAGACACATATTTCCC 1380
 1411 GTGCTTCAAGCTCTTTCTTCTGCTGCTGGAGTACAGCAGTGAAGCAAGACACATATTTCCC 1470
 Dh

Qy	1381	TGCTCAGAGATGACAAATGGCATCCTTGAGCTGACAAACATTTTCATGACCATAGG	1440
Db	1471	TGGTCACAGAGATGACAAATGGCATCCTTGAGCTGACAAACATTTTCATGACCATAGG	1530

QY	1441	TCACGTGTCTACACTGGGTACACTTTGTACGAGTGTGGCCCTCCACTGATGCTGTGCTCA	1500
Db	1531	TCACGTGTCTACACTGGGTACACTTTGTACGAGTGTGGCCCTCCACTGATGCTGTGCTCA	1590

QY 1501 GGACCTCTGTCCAGACAATCCCTTTACAAACAACACGCTGCTTGATCTTGTA 1560

Db 1591 GGACCTCTGTCCAGACAATCCCTTTACAAACAACACGCTGCTTGATCTTGTA 1650

QY	1561	CCCTTTACAGAGAAAGGAGGATCCCTGTGCCAAAGCTCCAGGCTCTCCCTGCAACT	1620
Db	1651	CCCTTTACAGAGAAAGGAGGATCCCTGTGCCAAAGCTCCAGGCTCTCCCTGCAACT	1710

QY	1622	CAGAGCCAGCCAGCTCAGCTGGGA	CTGTGTTCCAGACATCTGTCTCTTGAAT	1680
Db	1711	CAGAGCCAGCCAGCTCAGCTGGGA	CTGTGTTCCAGACATCTGTCTCTTAAT	1770

Qy	1681	AAGGATTCTCCTTCAGGCGTAAGCCGAGATTGGGCCAAGATTAAGATCCAACTA	1740
Db	1771	AAGGATTCTCCTTCAGGCGTAAGCCGAGATTGGGCCAAGATTAAGATCCAACTA	1830

Oy	1741	TGAGGCTGTTCTGTCTAACTCAAGCTGCTTGAA	TGAGGGTCCAGGCGCTGTCAAC	1800
Db	1831	TGAGGCTGTTCTGTCTAACTCAAGCTGCTTGAA	TGAGGGTCCAGGCGCTGTCAAC	1890

Oy 1801 ATGGGGCTTCTGACCTGACCAAGGTTGAGGACAGGATATGGCAGGGTCTGTCCGT 1860
 Db 1891 ATGGGGCTTCTGACCTGACCAAGGTTGAGGACAGGATATAGGCAGGGTCTGTCCGT 1950

Db
1851 GGCACCTGGAAAGTCCAGGTGGGACCTCTTCTGGGGACACTGGGGTCCACATCCAG 1920
1951 GGCACCTGGAAAGTCCAGGTGGGACCTCTTCTGGGGACACTGGGGTCCACATCCAG 2010

OY 1921 GTCCACTGAGTTTGGATACATGATGATGTTTACCTGTGCTAATAAGCA 1980
 DB 2011 GTCCACTGAGTTTGGATACATGATGATGTTTACCTGTGCTAATAAGCA 2070
 OY 1981 GAATTATGAAATATAAAAAA 2002
 DB 2071 GAATTATGAAATATAAAAAA 2092
 RESULT 4
 AAC77866
 ID AAC77866 standard, cDNA, 2048 BP.
 AC AAC77866;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:260.
 XX
 KM Human, cancer associated gene; cancer antigen; detection; cancer;
 KM diagnosis; cytotoxic; proliferative; vunerary; immunomodulator;
 KM antidiabetic; antiaesthetic; antirheumatic; anticholel; antiviral;
 KM antinflammatory; antihypertoid; antiallergic; antibacterial; cardiac;
 KM dermatological; neuroprotective; thrombolytic; coagulant; noctropic;
 KM vasotropic; antipneumatic; antidiabetic; gene therapy; inflammation;
 KM immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KM allergic reaction; graft versus host disease; organ rejection;
 KM haemostatic; thrombolytic; cardiovascular disorder; infection;
 KM neurological disease; drug screening; 88.
 XX
 OS Homo sapiens.
 XX
 PN WO20005350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PE 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI, 2000-587533/55.
 DR P-PSDB; AAB43657.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 1, Page 825-826; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytotoxic; proliferative; vunerary; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; anticholel; antiviral;
 CC antinflammatory; antihypertoid; antiallergic; antibacterial; cardiac;
 CC dermatological; neuroprotective; thrombolytic; coagulant;
 CC vasotropic; antipneumatic; antidiabetic; gene therapy; inflammation;
 CC immune disorder; haematopoietic cell disorder; autoimmune disorder;
 CC allergic reaction; graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.

XX
 SQ Sequence 2048 BP; 404 A; 619 C; 585 G; 434 T; 6 other;
 Query Match 85.9%; Score 1767.8; DB 21; Length 2048;
 Best Local Similarity 95.9%; Pred. No. 1.1e-303;
 Matches 1882; Conservative 4; Mismatches 1; Indels 76; Gaps 4;
 OY 49 ATGAGAGCCACCCCTCTGGCTGCTCTGCGGGTCTCTGTCAGAGAAAGCCGTGGAG 108
 DB 156 ATGAGAGCCACCCCTCTGGCTGCTCTGCGGGTCTCTGTCAGAGAAAGCCGTGGAG 215
 OY 109 TTGATGACAACTTATGATACCGAGGGTCCCGTCCAGAAAGAGCTGGAAGTGGCCCCAG 168
 DB 216 TTGATGACAACTTATGATACCGAGGGTCCCGTCCAGAAAGAGCTGGAAGTGGCCCCAG 275
 OY 169 CCCAGACTGCCCCCTGCTGTGGCCCTGAGCCCACTACTGCTCAGATCGTGCAACT 228
 DB 276 CCCAGACTGCCCCCTGCTGTGGCCCT----- 304
 OY 229 GCTGTGGCCACTGCTCTCCGCTTGGGCCCTATGTCCTCTGAGAGCCGAGAGAGCGGG 288
 DB 305 -----GAGCCGAGAGAGCGGG 322
 OY 289 CGGGGCTACCGGGCCCTGACATGCCCCACAGCACTGATATACCTGCAAGGTATACCC 348
 DB 323 CGGGGCTACCAAG3-CCTGACATGCCCTTACAGGCACTGATATACCTGCAAG-GTATACCC 380
 OY 349 GTCCAGAAAGCCCTGAGCGCTGCTGAGAGCCCTACGCGCGGCTGCCCGGCAAGCATGTG 408
 DB 381 GTCCAGAAAGCCCTGAGCGCTGCTGAGAGCCCTATGCGCGGCTGCCCGGCAAGCATGTG 440
 OY 409 GCTGCGGCCCACTGAGGCTCTGGCTGCTGATCCAGCTCTCTTCACTCGGACC 468
 DB 441 GCTGCGGCCCACTGAGGCTCTGGCTGCTGATCCAGCTCTCTTCACTCGGACC 500
 OY 469 CATGGGAGCATGACAGCGCTGGTGGAGCGCCACCGATCCCTGAGGCTGAGGCTGCC 528
 DB 501 CATGGGAGCATGACAGCGCTGGTGGAG-CGCCACCGATCCCTGAGGCTGAGGCTGCC 559
 OY 529 GTGCTCTTCCGCGAGATGCGCACCGCCCTGAGCGCATCTGTACACAGACAGGTGCTCTG 588
 DB 560 GTGCTCTTCCGCGAGATGCGCACCGCCCTGAGCGCATCTGTACACAGACAGGTGCTCTG 619
 OY 589 CGTATCTCAAGCTGTGTGCTTTGTCTTCTGCTGACCGTGAAGAGAAAGCTGTGCTG 648
 DB 620 CGTATCTCAAGCTGTGTGCTTTGTCTTCTGCTGACCGTGAAGAGAAAGCTGTGCTG 679
 OY 649 GAGAACTGAGAGAGCTCCGCGCTGCTGAGCTGGGCCAGATGATTTCCCTGTGGACAAGCAC 708
 DB 680 GAGAACTGAGAGAGCTCCGCGCTGCTGAGCTGGGCCAGATGATTTCCCTGTGGACAAGCAC 739
 OY 709 GCGTCCAGAGCTTACGTGGAGACTGAGATCTACAGCTCAAGGAGCTCATACTCGAGCAAG 768
 DB 740 GCGTCCAGAGCTTACGTGGAGACTGAGATCTACAGCTCAAGGAGCTCATACTCGAGCAAG 799
 OY 769 GAGCGGATGCTGAGAGCTGGAGCGTGGCGCTTTTACCATGCTGAGCGGCACTAACCC 828
 DB 800 GAGCGGATGCTGAGAGCTGGAGCGTGGCGCTTTTACCATGCTGAGCGGCACTAACCC 859
 OY 829 TTCCAGAGCTGGAGAGCTGCTGCTCTTCCGCAAGATCCGCGGAGCTTACAGCTTGG 888
 DB 860 TTCCAGAGCTGGAGAGCTGCTGCTCTTCCGCAAGATCCGCGGAGCTTACAGCTTGG 919
 OY 889 CCTGAGAGCTCTCGAGCCCTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
 DB 920 CCTGAGAGCTCTCGAGCCCTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979
 OY 949 GCTGAAAGCTTACAGAGCAAGGATCTCTCTGACAGCCCTGAGCTGCAAGAGCCGATG 1008
 DB 980 GCTGAAAGCTTACAGAGCAAGGATCTCTCTGACAGCCCTGAGCTGCAAGAGCCGATG 1039
 OY 1009 CCTTAAAGCCCAAGCCGATCTCATCTCTGAGAGGCTGCGAGGATGCTCTGATGAGACTG 1068

Dh	1040	CCCTTAGCTCCAAACCCGATCCCATCTCTGGAGAGCTGCCAAGTGTCCCTGATGGAATG	10399
Oy	1069	GGGCTGAGACGAAGCCAGGGAGAAAGAGGGAGACAGAGAAAGTGTTCTGTATGGCTTAGGAC	11288
Dh	1100	GGGCTGAGACGAAGCCAGGGAGAAAGAGGGAGACAGAGAAAGTGTTCTGTATGGCTTAGGAC	11599
Oy	1129	CACCTTACTTACACGCTCAGCTGGCCAAAGTGTGATTGGGGGTGAGCTCCAAAGCCTT	11888
Dh	1160	CACCTTACTTACACGCTCAGCTGGCCAAAGTGTGATTGGGGGTGAGCTCCAAAGCCTT	1219
Oy	1189	CTCCGTGCTCTGAACCTGAGCCAAACCTTCAGTGCCTTCAGAAAGGAGAAAGGACAGAGC	1248
Dh	1220	CTCCGTGCTCTGAACCTGAGCCAAACCTTCAGTGCCTTCAGAAAGGAGAAAGGACAGAGC	1279
Oy	1249	CTGTGTGAGTGTGTGTGTATCACATCTGCTTTGTTCACACATGCAAGTTCTGTCTTG	1308
Dh	1280	CTGTGTGAGTGTGTGTGTATCACATCTGCTTTGTTCACACATGCAAGTTCTGTCTTG	1339
Oy	1309	GATGCTTATCAGTGTCCAAAGCCCTGTCTTGTGTGTGGAGATACAGCAAGTATGACAAAGGA	1368
Dh	1340	GATGCTTATCAGTGTCCAAAGCCCTGTCTTGTGTGTGGAGATACAGCAAGTATGACAAAGGA	1399
Oy	1369	GACAATATTCCTGCTCAGAGATGACAAACTGGACTCTTGTAGCTGACAAACTTTTC	1428
Dh	1400	GACAATATTCCTGCTCAGAGATGACAAACTGGACTCTTGTAGCTGACAAACTTTTC	1459
Oy	1429	CATGACATAGAGTCACTGTCTTACATCTGGGTACACTTTGTACCAAGTGTGGGCTTCCACTGA	1488
Dh	1460	CATGACATAGAGTCACTGTCTTACATCTGGGTACACTTTGTACCAAGTGTGGGCTTCCACTGA	1519
Oy	1489	TGCTGTGTCTCAGGACCTCTGTCCAAAGACATGCCCTTACAAACAAACAGCTGCT	1548
Dh	1520	TGCTGTGTCTCAGGACCTCTGTCCAAAGACATGCCCTTACAAACAAACAGCTGCT	1579
Oy	1549	TTGTATCTTGTACCTTTTCAGAGAAAGGAGATCCCTGTGCCAAAGCTTCAGAGCCTC	1608
Dh	1580	TTGTATCTTGTACCTTTTCAGAGAAAGGAGATCCCTGTGCCAAAGCTTCAGAGCCTC	1639
Oy	1609	TCCCTGTGCACTCAGAGACCCAAAGCCAGCTCACTGTGGAACTGTGTTCCAGCACTCT	1668
Dh	1640	TCCCTGTGCACTCAGAGACCCAAAGCCAGCTCACTGTGGAACTGTGTTCCAGCACTCT	1699
Oy	1669	GTCCCTTGAATTAAGAGATTCCCTCTCCAGGCTTAAAGCTGTGGGATTTGGGCAAGATAA	1728
Dh	1700	GTCCCTTGAATTAAGAGATTCCCTCTCCAGGCTTAAAGCTGTGGGATTTGGGCAAGATAA	1759
Oy	1729	GAATTCGCAATATGAGAGCTAGTCTTGTCTTAACTCAGACTGTCTGGAATGAGGGTCCA	1788
Dh	1760	GAATTCGCAATATGAGAGCTAGTCTTGTCTTAACTCAGACTGTCTGGAATGAGGGTCCA	1819
Oy	1789	GGCTGTTCACCAATGGGGCTTCTGACTGAGACCAAGGTTGAGGAGACAGATTAAGCAG	1848
Dh	1820	GGCTGTTCACCAATGGGGCTTCTGACTGAGACCAAGGTTGAGGAGACAGATTAAGCAG	1879
Oy	1849	GGTCTGTCTGTGGCACTGTGGAAAGTCCAGGTGGGACTCTCTTGTGGGGAACATTGGGGT	1908
Dh	1880	GGTCTGTCTGTGGCACTGTGGAAAGTCCAGGTGGGACTCTCTTGTGGGGAACATTGGGGT	1939
Oy	1909	CCAGCAATCCAGGCTCCATCTCTAGGTTTGGATACCAATGAGTATGTATTTTACTGTG	1968
Dh	1940	CCAGCAATCCAGGCTCCATCTCTAGGTTTGGATACCAATGAGTATGTATTTTACTGTG	1999
Oy	1969	CTTATTAAGAGAAATTATGAATTAAGAAAAAAAAAAAAAAAAAAAAA 2011	
Dh	2000	CTTATTAAGAGAAATTATGAATTAAGAAAAAAAAAAAAAAAAAAAAA 2042	

RESULT	5
AAS06709	
ID	AAS06709 standard; cDNA; 1077 BP
XX	
AC	AAS06709,
XX	

DT	12-SEP-2001	(first entry)	
XX			
DB			Polynucleotide sequence encoding human protein kinase #9.
XX			
KM			Human; protein kinase; PTK; SYK; cancer; cardiovascular disease;
KM			metabolic disorder; immune related disease; neurological disorder;
KM			neurodegenerative disorder; inflammatory disorder; infectious disease;
XX			reproductive disorder; gene therapy; ss.
OS			Home sapiens.
XX			
PN	WO200138503-A2.		
XX			
PD	31-MAY-2001.		
XX			
PF	22-NOV-2000; 2000WO-US32085.		
XX			
PR	24-NOV-1999; 99US-0167482.		
XX			
PA	(SUGEN) SUGEN INC.		
XX			
PI	Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;		
PI	Flanagan P, Clary D;		
XX			
DR	WPI. 2001-343950/36.		
XX	P-ESDB; AAU03509.		
PT			Nucleic acids encoding human kinase polypeptides, useful for preventing
PT			diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT			neural-associated diseases, and microbial infections -
XX			
PS	Example 1; Figure 1; 43pp; English.		
XX			
CC	AA506701-AA506757 encode for novel human protein kinases #1-57. The		
CC	novel protein kinases have been identified as members of the tyrosine		
CC	or serine/threonine kinase (PTK and STK) families. The polynucleotides		
CC	encoding protein kinases and the polypeptides may be used in the		
CC	prevention, diagnosis and treatment of diseases associated with		
CC	inappropriate kinase expression. For example, they may be used to treat		
CC	cancers (especially cancers of hematopoietic origin), cardiovascular		
CC	disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),		
CC	immune related diseases (e.g. rheumatoid arthritis), neurological		
CC	disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.		
CC	Parkinson's disease), inflammatory disorders (e.g. asthma), infectious		
CC	disease (e.g. HIV) and reproductive disorders (e.g. infertility).		
CC	Additionally, polynucleotides encoding protein kinases may be		
CC	used for gene therapy and as DNA probes in diagnostic assays.		
CC	The protein kinase polypeptides may be used as antigens in the production		
CC	of antibodies against the protein kinases and in assays to identify		
CC	modulators of protein kinase expression and activity.		
XX			
SQ	Sequence 1077 BP; 174 A; 373 C; 325 G; 205 T; 0 other;		
	Query Match	52.2%;	Score 1073.8; DB 22; Length 1077;
	Best Local Similarity	99.8%;	Pred. No. 6e-181;
	Matches 1075; Conservative	0;	Mismatches 2; Indels 0; Gaps 0;
QY	49 ATGGGAGCCACCCCTTGCTGCTCTGCGGGTTCCTGTCCAGAGAGACGGTTGGAG	108	
DB	1 ATGGGAGCCACCCCTTGCTGCTCTGCGGGTTCCTGTCCAGAGAGACGGTTGGAG	60	
QY	109 TTGGATGCAACTTNGATACCGAGCGTCCGTCGAGAAAGAGCTCGAAGTGGGCCGAG	168	
DB	61 TTGGATGCAACTTNGATACCGAGCGTCCGTCGAGAAAGAGCTCGAAGTGGGCCGAG	120	
QY	169 CCCAGACTGCCCCCTCGCTGTTGCCCTCGAGCCCACTACTGTGCTCCAGATCGTCACT	228	
DB	121 CCCAGACTGCCCCCTCGCTGTTGCCCTCGAGCCCACTACTGTGCTCCAGATCGTCACT	180	
QY	229 GCTGTGGCACTGCTCCGCTTGGGCGCTATGCTCTCTGAGAGCCGAGAGAGGCGAG	288	
DB	181 GCTGTGGCACTGCTCCGCTTGGGCGCTATGCTCTCTGAGAGCCGAGAGAGGCGAG	240	


```

Db      181 GCTGAGGACATGAGCTCCGCTTGGGACCTATGCTCTGAGACCCGAGAGGCGGG 240
Qy      289 CGGGCTTACGGGGCCCTGCACTGCTTACAGGCACTGATATCTGCAAGGTGATACCC 348
Db      241 CGGGCTTACGGGGCCCTGCACTGCTTACAGGCACTGATATCTGCAAGGTGATACCC 300
Qy      349 GTCCAGAGAGGCGCTGAGCGGTGAGGCGCTTACGCGGCGGCTGCGCCGACAGAGATG 408
Db      301 GTCCAGAGAGGCGCTGAGCGGTGAGGCGCTTACGCGGCGGCTGCGCCGACAGATG 360
Qy      409 GCTGCGCCCACTGAGGTCTGAGCTGATACCAAGCTCTTACAGGCTTTTTCATCTGACC 468
Db      361 GCTGCGCCCACTGAGGTCTGAGCTGATACCAAGCTCTTACAGGCTTTTTCATCTGACC 420
Qy      469 CATGGGAGATGACACAGCTGTGTGAGAGCGCCACCGTATTCCTGAGGCTGAGGCTCC 528
Db      421 CATGGGAGATGACACAGCTGTGTGAGAGCGCCACCGTATTCCTGAGGCTGAGGCTCC 480
Qy      529 GTGCTCTTCCGCGAGATGAGCGCCCTGAGCGGCTGACACAGACAGGTCTGCTGCTG 588
Db      481 GTGCTCTTCCGCGAGATGAGCGCCCTGAGCGGCTGACACAGACAGGTCTGCTGCTG 540
Qy      589 CGTATCTCAAGCTGTGTGCTTGTCTTGC-----TGAACCTGAGAGAGAGAGCTG 642
Db      541 CGTATCTCAAGCTGTGTGCTTGTCTTGC-----TGAACCTGAGAGAGAGAGCTG 600
Qy      643 GTGCTGAGAGAGCTGAGAGAGCTCTGCTGTGCTGAGCTGAGGCGACAGATGATTCCTGTGGAC 702
Db      601 GTGCTGAGAGAGCTGAGAGAGCTCTGCTGTGCTGAGCTGAGGCGACAGATGATTCCTGTGGAC 660
Qy      703 AAGCAGCGTGTCCAGGCTCTAGTGGAGCTGAGATCTCAGCTCAGCGGCGCTCACTACCTG 762
Db      661 AAGCAGCGTGTCCAGGCTCTAGTGGAGCTGAGATCTCAGCTCAGCGGCGCTCACTACCTG 720
Qy      763 GGCAGAGCAGCGGATGTCTGAGAGCTGTGGCGCTTTCACCATGCTGAGCGGCGAC 822
Db      721 GGCAGAGCAGCGGATGTCTGAGAGCTGTGGCGCTTTCACCATGCTGAGCGGCGAC 780
Qy      823 TACCCCTTCCAGAGCTGAGAGCTGTGCTCTTCCGAGAGATTCGCGCGGCGCTTAC 882
Db      781 TACCCCTTCCAGAGCTGAGAGCTGTGCTCTTCCGAGAGATTCGCGCGGCGCTTAC 840
Qy      883 GCTTGTGCTGAGAGCTCTGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 942
Db      841 GCTTGTGCTGAGAGCTCTGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 900
Qy      943 GAGCAGCTGAGCGGCTCAGACGCAAGGATCTCTGCAACCCCTGCTGCAAGAGAC 1002
Db      901 GAGCAGCTGAGCGGCTCAGACGCAAGGATCTCTGCAACCCCTGCTGCAAGAGAC 960
Qy      1003 CCGATGCGCTTACGCGCCCAACCCGATCTCTCTGAGAGGTGCGCCAGAGTGTCTCTGAT 1062
Db      961 CCGATGCGCTTACGCGCCCAACCCGATCTCTCTGAGAGGTGCGCCAGAGTGTCTCTGAT 1020
Qy      1063 GGAAGTGGAGCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122
Db      1021 GGAAGTGGAGCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy      1123 TAG 1125
Db      1081 TAG 1083

```

RESULT 7

AA261155 standard; cDNA, 1085 BP.

AA261155;

30-MAY-2000 (first entry)

cDNA J5503-KS encoding domains VIA to XI of a protein kinase.

```

KW      Kinase activity; molecular weight marker; isoelectric focusing marker;
KM      peptide fragmentation control; cellular signal transduction; 88.
XX      Homo sapiens.
OS      Homo sapiens.
FH      Key
PT      CDS
FT      2..1081
       /*tag= a
       /note= "partial sequence"
XX      MO200008180-A2.
XX      17-FEB-2000.
XX      03-AUG-1999; 99WO-US17630.
XX      04-AUG-1998; 98US-0095270.
XX      11-SEP-1998; 98US-0099972.
XX      (IMMV ) IMMUNEX CORP.
XX      Virca GD, Bird TA, Anderson DM, Marken JS;
XX      WPI: 2000-195584/17.
XX      P-PSDB; AAY69157.
XX      New human kinase polypeptides and polynucleotides used as molecular
PT      weight markers and as controls for peptide fragmentation
XX      Claim 1; Page 7; 60pp; English.
XX      The present sequence encodes a partial polypeptide which has kinase
CC      activity. The kinase polynucleotides can be used to express the
CC      polypeptides, and as probes to identify nucleic acids encoding
CC      proteins having kinase activity. The kinase polypeptides and
CC      fragmented polypeptides are used as molecular weight and isoelectric
CC      focusing markers, and as controls for peptide fragmentation. They also
CC      have a number of therapeutic uses as kinases play a central role in
CC      cellular signal transduction. The polypeptides could also be used to
CC      identify binding partner proteins. The polypeptides can also be used as
CC      a reagent to identify any proteins that the polypeptide regulates, and
CC      proteins with which it might interact. The polypeptides may also be
CC      used for preparation of antibodies. The antibodies can be used in
CC      assays to detect the presence of the protein, and to purify the protein
CC      by immunoinfinity chromatography.
XX      Sequence 1085 BP; 164 A; 389 C; 338 G; 194 T; 0 other;
SQ      Query Match 48.8%; Score 1004.4; DB 21; Length 1085;
       Best Local Similarity 99.9%; Pred. No. 1.1e-168;
       Matches 1005; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1 GCTTGAAGCCCGGCGGCGCCGAGCCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db      80 GCTTGAAGCCCGGCGGCGCCGAGCCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 139
Qy      61 CTTTGAAGCTCTCTGAGGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db      140 CTTTGAAGCTCTCTGAGGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 199
Qy      121 TTAGATACCGAGAGTCCCGTCCAGAAACGAGCTGAGAGTGGGCGCCAGACAGAGAG 180
Db      200 TTAGATACCGAGAGTCCCGTCCAGAAACGAGCTGAGAGTGGGCGCCAGACAGAGAG 259
Qy      181 CCGTGCCTGTGGCCCTGAGAGCCAGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db      260 CCGTGCCTGTGGCCCTGAGAGCCAGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 319
Qy      241 GCTTCCGCTTGTGGGCTTATGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db      320 GCTTCCGCTTGTGGGCTTATGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
Qy      301 GCGCTGACAGTGCCTTACAGGCACTGAGTATCTCTGACAGAGTATACCCGCTCCAGAGAGCC 360

```



```

Db      380 GCGCTGCACTGCCCCCTCAAGAGCACTGAGTATACCTGCAAGAGTACCCGTCAGAGAACCC 439
Qy      361 CTGGCCGCTGTGAGACCCCTTAAGCGCGGCTGGCCCCCGCAAGACATGTGCTCGGCCACT 420
Db      440 CTGGCCGCTGTGAGACCCCTTAAGCGCGGCTGGCCCCCGCAAGACATGTGCTCGGCCACT 499
Qy      421 GAGGTCTGCTGCTGATACCAAGCTCTTAAGCTCTTTTCACTCGGACCCATGGGGACATG 480
Db      500 GAGGTCTGCTGCTGATACCAAGCTCTTAAGCTCTTTTCACTCGGACCCATGGGGACATG 559
Qy      481 CACAGCCTGTGTGCGAAGCGGACCGATATCCCTGAGCCTGAGAGCTGCGGTCTCTTCGCG 540
Db      560 CACAGCCTGTGTGCGAAGCGGACCGATATCCCTGAGCCTGAGAGCTGCGGTCTCTTCGCG 619
Qy      541 CAGATGCGCAACCGCCCTGAGCGCACTGTACAGCAACGAGCTGTGTCTTCTGCTGATCTCAAG 600
Db      620 CAGATGCGCAACCGCCCTGAGCGCACTGTACAGCAACGAGCTGTGTCTTCTGCTGATCTCAAG 679
Qy      601 CTGTGTGCTTTTGTCTTCTGCTGACCGGTGAGAGAAAGAGCTGTGTCTTCTGAGAACTTGAG 660
Db      680 CTGTGTGCTTTTGTCTTCTGCTGACCGGTGAGAGAAAGAGCTGTGTCTTCTGAGAACTTGAG 739
Qy      661 GACTCTGCGGTGCTGACTGTGGGCGAGATGATTCCTGTGTGGACAAGCAAGCGGTGCCAGCC 720
Db      740 GACTCTGCGGTGCTGACTGTGGGCGAGATGATTCCTGTGTGGACAAGCAAGCGGTGCCAGCC 799
Qy      721 TACGTGTGACCTGTGAGATATCAAGCTCAAGGCGCTTATATCTGGGCAAGCAAGCGGTATTC 780
Db      800 TACGTGTGACCTGTGAGATATCAAGCTCAAGGCGCTTATATCTGGGCAAGCAAGCGGTATTC 859
Qy      781 TGAAGCTGTGGCGGTGTGGCGCTTTCACCAATGTGTGGCGGCACTACCTTTCAGAGACTGCG 840
Db      860 TGAAGCTGTGGCGGTGTGGCGCTTTCACCAATGTGTGGCGGCACTACCTTTCAGAGACTGCG 919
Qy      841 GAGCTGTGCTGTCTTCTGAGCAAGATCCGCGCGGGGCGCTAGCGCTTCTGAGAGGCTTC 900
Db      920 GAGCTGTGCTGTCTTCTGAGCAAGATCCGCGCGGGGCGCTAGCGCTTCTGAGAGGCTTC 979
Qy      901 TCGGCCCCCTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Db      980 TCGGCCCCCTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1039
Qy      961 ACAGCCACAGGCAATCTCTGCAAGCCCGCTGGCTGGAGCAAGACCCGA 1006
Db      1040 ACAGCCACAGGCAATCTCTGCAAGCCCGCTGGCTGGAGCAAGACCCGA 1085

RESULT 8
AAH76218
ID      AAH76218 standard, cDNA, 972 BP.
XX
XX      AAH76218;
XX
XX      29-OCT-2001 (first entry)
XX
XX      Human kinase PKIN-10 encoding cDNA.
XX
XX      PKIN; kinase; cytosolic; immunosuppressive; immunostimulant; human;
XX      antiarteriosclerotic; cardiac; gene therapy; antisense therapy; 88.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..972
XX      PT      /*tag= a
XX      /product= "PKIN-10"
XX
XX      MO200160991-A2.
XX
XX      23-AUG-2001.
XX
XX      16-FEB-2001; 2001WO-US05240.
PP

```

```

XX      17-FEB-2000; 2000US-0183682.
PR      02-MAR-2000; 2000US-0186559.
PR      09-MAR-2000; 2000US-0188606.
PR      17-MAR-2000; 2000US-0189998.
PR      30-MAR-2000; 2000US-0193851.
XX
XX      (INCY-) INCYTE GENOMICS INC.
XX
XX      Tang YT, Buford N, Gandhi AR, Patterson C, Khan PA, Yue H,
XX      Hefajia A, Shih LL, Tribouley CM, Yao MG, Burrill JD, Marcus GA,
XX      Zingler RA, Lu DM, Bandman O, Policky JL, Griffin JA, Thornton M,
XX      Nguyen DB, Lai P, Walsh RT;
XX
XX      WPI; 2001-514771/56.
XX
XX      P-PSDB; AAB85791.
XX
XX      Isolated human kinase polypeptides useful in the diagnosis, treatment
XX      and prevention of cancer, immune disorders and disorders affecting
XX      growth and development -
XX
XX      Claim 5; Page 125; 126pp; English.
XX
XX      The invention provides human kinases (PKIN) and polynucleotides encoding
XX      PKIN. The PKIN polypeptides can be expressed using standard recombinant
XX      methodology. The PKIN polypeptides, polynucleotides, modulators and
XX      specific antibodies are useful in the diagnosis, treatment and prevention
XX      of cancer, immune disorders, disorders affecting growth and development,
XX      atherosclerosis, and other cardiovascular diseases, and lipid disorders
XX      and in the assessment of the effects of exogenous compounds on the
XX      expression of nucleic acid sequences of human kinases. The present
XX      sequence represents a cDNA encoding a human PKIN-10 polypeptide.
XX
XX      Sequence 972 BP; 167 A; 329 C; 294 G; 182 T; 0 other;
XX
XX      Query Match      39.0%; Score 802.2; DB 22; Length 972;
XX      Best Local Similarity 87.7%; Pred. No. 6.7e-133;
XX      Matches 944; Conservative 0; Mismatches 28; Indels 105; Gaps 2;
XX
Qy      49 ATGCGAGCCACCCCTGTGCTGCTCTGCGGCTTCCCTGTCCAGAGAAAGCGTTTGAG 108
Db      1 ATGCGAGCCACCCCTTTGGCTGCTTGGCGGCTTCCCTGTCCAGAGAAAGCGTTTGAG 60
Qy      109 TTGATATGCAACTTATATATACGAGCGTCCGCTCCAGAAAGAGCTTCAGAGTGGCCCGAG 168
Db      61 TTGATATGCAACTTATATATACGAGCGTCCGCTCCAGAAAGAGCTTCAGAGTGGCCCGAG 120
Qy      169 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 228
Db      121 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Qy      229 GCTGTGCACTGCTCCGCTTGTGGCCCTATGCTCTCTGAGAGCCGAGAGAGGCGGG 288
Db      181 GCTGTGCACTGCTCCGCTTGTGGCCCTATGCTCTCTGAGAGCCGAGAGAGGCGGG 240
Qy      289 CGGACTTACCGGCGCTGTGCACTGCTTACAGGCACTAGATATATCTGAGAGGTATACCC 348
Db      241 CGGACTTACCGGCGCTGTGCACTGCTTACAGGCACTAGATATATCTGAGAGGTATACCC 300
Qy      349 GTCCAGAGAGCGCTGCGGCTGTGAGAGCCCTACGCGGCGCTGCGCCGCAAGACATGTG 408
Db      301 GTCCAGAGAGCGCTGCGGCTGTGAGAGCCCTATGCGGCGCTGCGCCGCAAGACATGTG 360
Qy      409 GCTGAGCCCACTGAGTCTGTGCTGTGTATCCAGTCTCTTACAGCTTTTTCATCTGAGCC 468
Db      361 GCTGAGCCCACTGAGGCTGTGCTGTGTATCCAACTTCTTACAGCTTTTTCATCTGAGCC 420
Qy      469 CATGGGAGCATGCAAGCTGTGTGCGAAGCGGACCGATATCCCTGAGAGCTGAGGCTGCC 528
Db      421 CATGGGAGCATGCAAGCTGTGTGCGAAGCGGACCGATATCCCTGAGAGCTGAGGCTGCC 486
Qy      529 GTGCTTTTCCGCGAGATGCGACCGCTGCGCACTGTGACAGCAAGGATGTGTCTGTG 588

```


Db 447 -----GCAACGCCCTGTCGCACTGTGACGAG----- 474
 QY 589 CGTATCTCAAGCTGTGTCTTTGTTCTGTGTCGACCGTGTAGAGAAAGAGTGTCTGT 648
 Db 475 -----ACACGGAAGAGCTGTGTCTGT 495
 QY 649 GAGAACTGTGAGACTCTGT 708
 Db 496 GAGAACTGTGAGACTCTGT 555
 QY 709 GCGTGTGAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGT 768
 Db 556 GCGTGTGAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGT 615
 QY 769 GCAAGCCGATGTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 828
 Db 616 GCAAGCCGATGTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 675
 QY 829 TTTCAGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 888
 Db 676 TTTCAGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 735
 QY 889 CCTGACAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGT 948
 Db 736 CCTGACAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 795
 QY 949 GCTGAAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 1008
 Db 796 GCTGAAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 855
 QY 1009 CCTTGAAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 1068
 Db 856 CCTTGAAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACT 915
 QY 1069 GGGCTGACGAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAG 1125
 Db 916 GGGCTGACGAGCTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAGACTGTGAG 972

RESULT 9
 ABL39747
 ID ABL39747 strand; cDNA, 1076 BP.
 XX ABL39747;
 AC
 XX
 DT 10-MAY-2002 (first entry)
 XX
 DE Human NS cDNA sequence SEQ ID NO:57.
 XX
 KW Human, cytostatic, osteopathic, gynaecological, neuroprotective;
 KW antirheumatic, antiarthritic, antipeptic, ophthalmological, anti-HIV;
 KW vasorelaxant, antidiabetic, antihypertensive, antineoplastic, dermatologic;
 KW anorectic, muscular, antileukemic, cardiovascular, anticonvulsant;
 KW anticonvulsant, hypotensive, antileukemic, immunomodulatory, cardiac;
 KW anticonvulsant, antidiabetic, tranquilizer, antidepressant, neuroleptic;
 KW gastroenteric, vitamin, antileukemic, antidiabetic, neuroprotective, neurotropic;
 KW contractile, vaccine, gene therapy, cancer, osteoporosis, dystonia;
 KW endometrial, degenerative disease, multiple sclerosis, psoriasis;
 KW rheumatoid arthritis, catarrh, xeroderma, atherosclerosis, glaucoma;
 KW inflammation, skin disorder, obesity, muscular dystrophy, AIDS;
 KW infertility, cardiovascular disease, coagulation disease, hypertension;
 KW leukaemia, asthma, immune disease, epilepsy, angina, neurodegeneration;
 KW diabetes, anxiety, depression, schizophrenia, viral disease, stroke;
 KW gastric ulcer, Alzheimer's disease, gene, ss.
 KW
 OS Homo sapiens.
 XX
 PN MO200206315-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001, 2001WO-1100653.
 XX

PR 18-JUL-2000; 2000IL-0137345.
 PR 15-DEC-2000; 2000IL-0140354.
 PA (COMP-) COMPUGEN LTD.
 PI Mintz L, Freilich S, Bernstein J;
 DR WPI; 2002-155037/20.
 DR P-PSDB; ABB06093.
 PT One hundred and twenty eight novel nucleic acid sequences, useful for
 XX treating and diagnosing e.g. cancer, asthma and Alzheimer's -
 PS Claim 1, Page 107; 290pp, English.
 XX
 CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antirheumatic, antiarthritic, antipeptic, ophthalmological, vitaminic,
 CC vasorelaxant, antidiabetic, antihypertensive, antineoplastic, dermatologic,
 CC anorectic, muscular, antileukemic, hypotensive, cardiovascular, cardiac,
 CC anticonvulsant, antidiabetic, tranquilizer, antidepressant, neuroleptic,
 CC immunomodulatory, antidiabetic, antileukemic, immunomodulatory, cardiac,
 CC antidepressant, gastroenteric, antidiabetic, neuroleptic, neuroprotective,
 CC neurotropic and contractile activities. The NS can be used in vaccines,
 CC gene therapy and antileukemic therapy. Nucleic acid, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometrial, psoriasis, atherosclerosis,
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC catarrh, xeroderma, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, leukaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contractile.
 XX
 SQ Sequence 1076 BP; 172 A; 352 C; 346 G; 194 T; 12 other;

Query Match 31.3%; Score 645.2; DB 24; Length 1076;
 Best Local Similarity 98.2%; Pred. No. 4e-105;
 Matches 670; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

QY 1 GCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 60
 Db 281 GCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 340
 QY 61 CCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 120
 Db 341 CCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 400
 QY 121 TTGATATACGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 180
 Db 401 TTGATATACGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 460
 QY 181 CCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 240
 Db 461 CCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 520
 QY 241 GCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 300
 Db 521 GCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 580
 QY 301 GCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 360
 Db 581 GCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 640
 QY 361 CTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 420
 Db 641 CTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 700
 QY 421 GAGGTCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 480
 Db 701 GAGGTCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 760


```

Db      1518 GTGGAGCTGAGATACCTCAGCTCAGCGGCTCATATCTGGGCAAGGCAAGCGATGCTGG
      1459
Oy      784 AGCTGGGCGTGGCGCTCTTCAACCATGTGGCCGCACTACCCCTTCAGAGACTCGGAG
      843
Db      1458 AGCTGGGCGTGGCGCTCTTCAACCATGTGGCCGCACTACCCCTTCAGAGACTCGGAG
      1399
Oy      844 CTTGTCTGCTCTTGGCAAGATCCGCGCGGCGCTTACCGCTTGTGCAAGCCTTTCG
      903
Db      1398 CTTGTCTGCTCTTGGCAAGATCCGCGCGGCGCTTACCGCTTGTGCAAGCCTTTCG
      1339
Oy      904 GCGCCCTGCGGCTGTGTGTGCTGCTCTTGTGTGGGAGCGAGCTGAACGCTCACA
      963
Db      1338 CCGTCGCGGCGTGTGTGTGCTGCTCTTGTGTGGGAGCGAGCTGAACGCTCACA
      1279
Oy      964 GCCAAGAGCATCTCTGCAACCCCTGCTGCAAGAGACCCGATGCCCTTAAAGCCCAAC
      1023
Db      1278 G-CAGAGCATCTCTCTGCAACCCCTGCTGCAAGAGACCCGATGCCCTTAAAGCCCAAC
      1221
Oy      1024 CGATCCCATCTTGGAGAGGCTGCCAGGTGTCTCTGATGAGACTGGGGCTGGAACGAA
      1083
Db      1220 CGATCCCATCTTGGAGAGGCTGCCAGGTGTCTCTGATGAGACTGGGGCTGGAACGAA
      1162
Oy      1084 AGGGAAGAGAGGAGAGAGAGAGAGAGTGTCTGTATGAGCTAGAGCAACCCCTACTAGAC
      1143
Db      1161 AGGGAAGAGAGGAGAGAGAGAGAGAGTGTCTGTATGAGCTAGAGCAACCCCTACTAGAC
      1102
Oy      1144 TCAGCTGCCAAGATGATGATTTGGGGGTAGCTCAAGCCCTTCTGCTGCTCTGA
      1201
Db      1101 TCAGCTGCCAAGATGATGATTTGGGGGTAGCTCAAGCCCTTCTGCTGCTCTGA
      1044

RESULT 14
ABL87304/c
ID      ABL87304 standard; cDNA; 426 BP.
AC      ABL87304;
XX
XX      17-MAY-2002 (first entry)
XX
XX      Human ovarian cancer related cDNA clone SEQ ID NO:10282.
XX
XX      Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
XX      Homo sapiens.
XX
XX      WO200192581-A2.
XX
XX      PD      06-DEC-2001.
XX
XX      PP      29-MAY-2001; 2001WO-US17756.
XX
XX      PR      26-MAY-2000; 2000US-207484P.
XX
XX      PA      (CORI-) CORIXA CORP.
XX
XX      PI      Algate PA, Harlocker SL, Jones R;
XX
XX      DR      WPI; 2002-122075/16.
XX
XX      PT      Composition for therapy and diagnosis of ovarian cancer comprising
XX      polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
XX      polypeptide, antibody specific to polypeptide or T cell expressing
XX      polypeptide
XX
XX      PS      Claim 1; SEQ ID 10282; 489bp; English.
XX
XX      The present invention describes a composition (I) comprising: carriers
XX      and immunostimulants; and a polypeptide (II) of a ovarian tumour
XX      polypeptide encoded by a polynucleotide (III) having a cDNA sequence
XX      (S1) from the 10912 nucleotide sequences as given in ABL77023 to
XX      ABL87934, (III) encoding (II) having a sequence (S2), a T cell
XX      population of (II), or antigen presenting cells that express (II).

```

```

CC      (I) has cytostatic activity. An oligonucleotide (IV) that hybridizes to
CC      (S1) can be used for detecting ovarian cancer in a patient's biological
CC      sample preferably serum or ovarian tissue. The method comprises
CC      contacting a biological sample from a patient with (IV), detecting the
CC      amount of polynucleotide hybridising to (IV) and comparing the amount to
CC      a predetermined cutoff value and thereby detecting ovarian cancer in the
CC      patient, where the amount of polynucleotide hybridising to (IV) is
CC      detected preferably by polymerase chain reaction (PCR). (I) comprising
CC      (III) and/or (II) is useful for stimulating and/or expanding T cells
CC      specific for an ovarian tumour protein comprising contacting T cells
CC      with (III) or (II). (III) is useful in design and preparation of
CC      ribozyme molecules for inhibiting expression of the tumour polypeptides
CC      and proteins in tumour cells; and to isolate a full length gene from a
CC      suitable library e.g., a tumour cDNA library using well known
CC      techniques.
XX
XX      SQ      Sequence 426 BP; 105 A; 112 C; 107 G; 102 T; 0 other;
XX
XX      Query Match      20.7%; Score 426; DB 24; Length 426;
XX      Best Local Similarity 100.0%; Pred. No. 2.1e-66;
XX      Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Oy      1567 CAGAGAAAGGAGATACCTGTCGCAAGGCTCCAGGCTTCCCTGCAACTCAGAGAC
      1626
Db      426 CAGAGAAAGGAGATACCTGTCGCAAGGCTCCAGGCTTCCCTGCAACTCAGAGAC
      367
Oy      1627 CCAAGCCAGCTCACTGTGGAACCTGTTCGCCAGCATCTGTCTCTTGAATTAAGAGA
      1686
Db      366 CCAAGCCAGCTCACTGTGGAACCTGTTCGCCAGCATCTGTCTCTTGAATTAAGAGA
      307
Oy      1687 TTCTCCTTCAAGGCTTAAGCTGGGATTTGGGGCAGAGATTAAGATCAACTTGAAGGC
      1746
Db      306 TTCTCCTTCAAGGCTTAAGCTGGGATTTGGGGCAGAGATTAAGATCAACTTGAAGGC
      247
Oy      1747 TAGTTCTTGTCTTAACCTCAAGACTGTCTGGAATGAGGCTCCAGGCTGTCAACCATGGG
      1806
Db      246 TAGTTCTTGTCTTAACCTCAAGACTGTCTGGAATGAGGCTCCAGGCTGTCAACCATGGG
      187
Oy      1807 CTTCTGACTGAGACCAAGGTTGAGGAGACAGATTAGGCAAGGCTGTCTGTGGCCAC
      1866
Db      186 CTTCTGACTGAGACCAAGGTTGAGGAGACAGATTAGGCAAGGCTGTCTGTGGCCAC
      127
Oy      1867 CTGGAAGTCCCAAGTGGGACTCTTCTGGGAGACCTTGGGGTCCACATATCCAGATCCAT
      1926
Db      126 CTGGAAGTCCCAAGTGGGACTCTTCTGGGAGACCTTGGGGTCCACATATCCAGATCCAT
      67
Oy      1927 ACTCTAGTTTGGATTCAGATGATGATGTTTACCTGAGCCTAATTAAGAGAAATTA
      1986
Db      66 ACTCTAGTTTGGATTCAGATGATGATGTTTACCTGAGCCTAATTAAGAGAAATTA
      7
Oy      1987 TGAAT 1992
      6 TGAAT 1
Db

RESULT 15
AAS37461
ID      AAS37461 standard; cDNA; 396 BP.
AC      AAS37461;
XX
XX      17-DEC-2001 (first entry)
XX
XX      Novel human diagnostic and therapeutic gene #519.
XX
XX      Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
XX
XX      Homo sapiens.
XX
XX      WO20016753-A2.
XX
XX      PD      13-SEP-2001.
XX

```

PF 09-MAR-2001; 2001MO-US07787.
XX
PR 09-MAR-2000; 2000US-0188609.
XX
PA (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kaessm A, Lamson G;
PI Drmanac R, Crkvenjakov R, Dickson M, Dickson S, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX
DR WPI; 2001-530177/58.
XX
XX New polynucleotides and polypeptides, useful for diagnosis and
PT treatment of breast, lung and colon cancer -
XX
XX
PS Claim 1; Page 727; 1193jp; English.
XX
CC The invention relates to new polynucleotides and polypeptides, useful for
CC diagnosis and treatment of breast, lung and colon cancer. The sequences
CC can be used in detecting differentially expressed genes correlated with a
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample derived from a
CC cell suspected of being cancerous. They can also be used to inhibit
CC tumour growth by modulating expression of a gene product. AAS35943-
CC AAS35938 represent novel human diagnostic and therapeutic coding
CC sequences of the invention.
XX
SQ Sequence 396 BP; 64 A; 138 C; 124 G; 70 T; 0 other;

Query Match 16.0%; Score 330.2; DB 22; Length 396;
Best Local Similarity 97.7%; Pred. No. 1.9e-49;
Matches 335; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCTCTGAGCCCGCGGCGCCCGGCGCCAGCGGAGCGAGATCGAGCCACC 60
DB 53 GCTCTGAGCCCGCGGCGCCCGGCGCCAGCGGAGCGAGATCGAGCCACC 112

QY 61 CCTCTGAGCTGCTCTGCGGGTTCCCTGTCCAGAAAGCGGTTGAGTTGATGACAC 120
DB 113 CCTCTGAGCTGCTCTGCGGGTTCCCTGTCCAGAAAGCGGATGAGTTGATGACAC 172

QY 121 TTAGATACCGAGGCTCCCGTCCAGAAAGAGCTGAAAGTGGGCCCGAGCCAGACTGGCC 180
DB 173 TTAGATACCGAGGCTCCCGTCCAGAAAGAGCTGAAAGTGGGCCCGAGCCAGACTGGCC 232

QY 181 CCTGAGCTGTTGCCCTGAGGCCACTAGTCTCAGATCGTGCACCTGCTGTGGCACT 240
DB 233 CCTGAGCTGTTGCCCTGAGGCCACTAGTCTCAGATCGTGCACCTGCTGTGGCACT 292

QY 241 GCTTCCCGCTTTGGGCGCTATGTCTCTCTGAGAGCCGAGAGAGGCGGCGGCTTACCG 300
DB 293 GCTTCCCGCTTTGGGCGCTATGTCTCTCTGAGAGCCGAGAGAGGCGGCGGCTTACCG 352

QY 301 GCCCTGCACTGCCCTACAGGCACTAGATATACCTGCAAGGTGT 343
DB 353 GCCCTGCACTGCCCTACAGGCACTAGATATACCTGCAAGGTGT 395

Search completed: January 16, 2004, 11:32:55
Job time : 595.591 secs